UNIVERSAL
LIBRARY

OU_172034

UNIVERSAL
LIBRARY

PHILOSOPHICAL

TRANSACTIONS

OF THE

ROYAL SOCIETY OF LONDO

SERIES A.

CONTAINING PAPERS OF A MATHEMATICAL OR PHYSICAL CHARACTE

VOL. 220.

PRINTED AND PUBLISHED FOR THE ROYAL SOCIETY BY
HARRISON AND SONS, 44–47, ST. MARTIN'S LANE, LONDON, W.C. 2,
Printers in Ordinary to His Majesty.

JUNE, 1920.

CONTENTS.

(A)

VOL. 220.

List of Illustrations . . . . page v
Advertisement vii

I. *Tidal Friction in the Irish Sea. By* G. I. TAYLOR, *M.A. Communicated by Sir* NAPIER SHAW, *F.R.S.* . . . . *page* 1

II. *Electromagnetic Integrals. By Sir* G. GREENHILL, *F.R.S.* 35

III. BAKERIAN LECTURE, 1918.—*Experiments on the Artificial Production of Diamond. By the Hon. Sir* CHARLES ALGERNON PARSONS, *K.C.B., F.R.S.* 67

IV. *The Pressure upon the Poles of the Electric Arc. By* W. G. DUFFIELD, *D.Sc., Professor of Physics, Captain, Royal Air Force,* THOS. H. BURNHAM, *B.Sc., and* A. H. DAVIS, *B.Sc., Post-Graduate Students in Physics, University College, Reading. Communicated by Prof.* O. W. RICHARDSON, *F.R.S.* . 109

V. *On Intensity Relations in the Spectrum of Helium. By* T. R. MERTON, *D.Sc., Lecturer in Spectroscopy, University of London, King's College, and* J. W. NICHOLSON, *F.R.S., Professor of Mathematics in the University of London* 137

VI. *The Scattering of Plane Electric Waves by Spheres. By* T. J. I'A. BROMWICH, *Sc.D., F.R.S.* 175

VII. *On the Fundamental Formulations of Electrodynamics. By* G. H. Livens, *University of Manchester. Communicated by* W. M. Hicks, *F.R.S.* *page* 207

VIII. *The Influence of Molecular Constitution and Temperature on Magnetic Susceptibility. Part IV.—Further Applications of the Molecular Field. By* A. E. Oxley, *Major, R.A.F., M.A., D.Sc., Mackinnon Student of the Royal Society. Communicated by Prof.* J. W. Nicholson, *F.R.S.* 247

IX. *A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations made at the Total Eclipse of May* 29, 1919. *By Sir* F. W. Dyson, *F.R.S., Astronomer Royal, Prof.* A. S. Eddington, *F.R.S., and Mr.* C. Davidson. (*Communicated by the Joint Permanent Eclipse Committee*) 291

X. *A Critical Study of Spectral Series. Part V.—The Spectra of the Monatomic Gases. By* W. M. Hicks, *F.R.S.* 335

Index to Volume 469

LIST OF ILLUSTRATIONS.

Plate 1.—Sir F. W. Dyson, Prof. A. S. Eddington, and Mr. C. Davidson on a Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations made at the Total Eclipse of May 29, 1919.

Plates 2–5.—W. M. Hicks: A Critical Study of Spectral Series. Part V.—The Spectra of the Monatomic Gases.

ADVERTISEMENT.

THE Committee appointed by the *Royal Society* to direct the publication of the *Philosophical Transactions* take this opportunity to acquaint the public that it fully appears, as well from the Council-books and Journals of the Society as from repeated declarations which have been made in several former *Transactions*, that the printing of them was always, from time to time, the single act of the respective Secretaries till the Forty-seventh Volume; the Society, as a Body, never interesting themselves any further in their publication than by occasionally recommending the revival of them to some of their Secretaries, when, from the particular circumstances of their affairs, the *Transactions* had happened for any length of time to be intermitted. And this seems principally to have been done with a view to satisfy the public that their usual meetings were then continued, for the improvement of knowledge and benefit of mankind: the great ends of their first institution by the Royal Charters, and which they have ever since steadily pursued.

But the Society being of late years greatly enlarged, and their communications more numerous, it was thought advisable that a Committee of their members should be appointed to reconsider the papers read before them, and select out of them such as they should judge most proper for publication in the future *Transactions;* which was accordingly done upon the 26th of March, 1752. And the grounds of their choice are, and will continue to be, the importance and singularity of the subjects, or the advantageous manner of treating them: without pretending to answer for the certainty of the facts, or propriety of the reasonings contained in the several papers so published, which must still rest on the credit or judgment of their respective authors.

It is likewise necessary on this occasion to remark, that it is an established rule of the Society, to which they will always adhere, never to give their opinion, as a Body, upon any subject, either of Nature or Art, that comes before them. And therefore the thanks, which are frequently proposed from the Chair, to be given to the authors of such papers as are read at their accustomed meetings, or to the persons through whose hands they received them, are to be considered in no other light than as a matter of civility, in return for the respect shown to the Society by those communications. The like also is to be said with regard to the several projects, inventions, and curiosities of various kinds, which are often exhibited to the Society; the authors whereof, or those who exhibit them, frequently take the liberty to report, and even to certify in the public newspapers, that they have met with the highest applause and approbation. And therefore it is hoped that no regard will hereafter be paid to such reports and public notices; which in some instances have been too lightly credited, to the dishonour of the Society.

VI. *The Scattering of Plane Electric Waves by Spheres.*

By T. J. I'A. BROMWICH, *Sc.D.*, *F.R.S.*

Received April 13,—Read November 23, 1916.

INTRODUCTORY NOTE.

THE problem which gives its title to the present paper has been handled by various writers, notably by Lord RAYLEIGH, Sir J. J. THOMSON, and Prof. LOVE. In most cases the solutions have been expressed in a Cartesian form; but it appears to me that a marked simplification is introduced by using spherical polar co-ordinates. The preliminary analysis becomes shorter, and the conclusions are easier to interpret; in fact, the analysis is nearly as simple as in the analogous problem of electrostatics, when an electric field is disturbed by the presence of a dielectric sphere.

To obtain the requisite solutions a new general solution of the electromagnetic equations in Cartesian form is given in § 1, and is then transformed to the spherical polar form; §§ 2, 3 contain a summary of certain analytical results required in the sequel.

§ 4 contains the general solution of the problem of finding the scattered waves when a plane simple harmonic wave strikes a sphere; and in § 5 the solution is applied to the case of a small sphere. These formulæ (all of § 4 and part of § 5) were originally worked out in 1899, but publication was postponed in the hope of completing the problem of the large sphere.

In § 6 the problem of a large sphere is considered by applying to the formulæ of § 4 a method of approximation devised by Prof. H. M. MACDONALD* for dealing with waves incident from a Hertzian oscillator on a conducting sphere. The formulæ of § 6 were worked out early in 1910 and were given in my University lectures at Cambridge in that year.†

At the same time I succeeded in obtaining a different treatment (given in § 7 below) which confirmed the other results, and gave an easier process for dealing with

---

* 'Phil. Trans. Roy. Soc.,' A, vol. 210, 1910, p. 113. Prof. MACDONALD tells me that he had worked out (at about the same time) results in reference to the problem of § 6; but these have not been published.

† An alternative solution was obtained by Prof. J. W. NICHOLSON at about the same time; his solution starts from Sir J. J. THOMSON'S formulæ. Prof. NICHOLSON'S results originally differed from those of § 6; but on revision agreement was obtained ('Proc. Lond. Math. Soc.,' vol. 9, 1910, p. 67; vol. 11, 1912, p. 277).

[Published February 2, 1920.

points behind the sphere. The method of this section is similar in some respects to one used by Prof. MACDONALD in a later paper.[*]

The formulæ of §§ 6, 7 have been delayed in publication for two reasons: in the first place I wished to obtain some confirmation from direct numerical calculation. This has now been carried out by Messrs. PROUDMAN, DOODSON and KENNEDY (of Liverpool University).[†] It appears that the agreement with the formulæ of § 6 is quite close (for $\kappa a = 9, 10$) from $\theta = 0°$ to $90°$, and for the Z-component up to about 120°. The formulæ of § 7 also give good results in a cone of about 10° behind the sphere (that is, from $\theta = 170°$ to $180°$). It is clear, however, that an approximation suitable from $\theta = 90°$ to $170°$ (for Y) and from $\theta = 120°$ to $170°$ (for Z) has still to be obtained. But nevertheless the present approximations proved a valuable auxiliary[‡] in checking and testing the numerical work.

[The paper in its original form was presented to the Society on April 13, 1916; owing to the difficulties in regard to labour and paper during the war, I was asked to condense the introductory matter of §§ 1–3. This proved to be impossible until now, on account of pressure of war-work of various kinds. In the present version § 1 has been re-written so as to reduce its bulk; in §§ 2, 3 certain formulæ have been omitted which were not used in the applications of §§ 4–6.

In re-arranging the paper it proved convenient also to number the formulæ differently. The decimal system has now been adopted; here the figure before the decimal point indicates the section of the paper in which the formula occurs. The figures following the decimal point are to be regarded as following the same order as ordinary decimal fractions. Thus (5·21) and (5·22) fall between (5·2) and (5·3), and all these formulæ occur in § 5.—*Added March* 18, 1919.]

§ 1. A GENERAL SOLUTION OF THE FUNDAMENTAL ELECTROMAGNETIC EQUATIONS.[§]

The fundamental equations of electromagnetic waves may be written

$$(E)\qquad \frac{K}{c^2}\frac{\partial X}{\partial t} = \frac{\partial \gamma}{\partial y} - \frac{\partial \beta}{\partial z}, \qquad \frac{K}{c^2}\frac{\partial Y}{\partial t} = \frac{\partial \alpha}{\partial z} - \frac{\partial \gamma}{\partial x}, \qquad \frac{K}{c^2}\frac{\partial z}{\partial t} = \frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y}.$$

$$(M)\qquad -\mu\frac{\partial \alpha}{\partial t} = \frac{\partial Z}{\partial y} - \frac{\partial Y}{\partial z}, \qquad -\mu\frac{\partial \beta}{\partial t} = \frac{\partial X}{\partial z} - \frac{\partial Z}{\partial x}, \qquad -\mu\frac{\partial \gamma}{\partial t} = \frac{\partial Y}{\partial x} - \frac{\partial X}{\partial y}.$$

---

[*] 'Phil. Trans. Roy. Soc.,' A, vol. 212, 1912, p. 299. The two methods are not identical; but they appear to yield equivalent results in all the cases to which they have been applied.

[†] 'Phil. Trans. Roy. Soc.,' A, vol. 217, 1917, p. 279. The calculation was originally undertaken by Dr. PROUDMAN in consequence of a suggestion made in my lectures of 1912; the work, however, proved to be longer than had been anticipated and was completed by Messrs. DOODSON and KENNEDY.

[‡] See the paper last quoted, p. 292 *et seq.*

[§] Revised March 18, 1919; see note at the end of the introductory remarks above

Here (X, Y, Z) denotes the electric force, (α, β, γ) the magnetic force, K is the dielectric constant, μ is the magnetic permeability, and the axes of reference are a Cartesian right-handed system. The units adopted are those of the electromagnetic system, and c is the fundamental constant generally identified with the velocity of radiation in free space; the equations (E) are those derived from AMPÈRE'S law, and the equations (M) are similarly derived from FARADAY'S law, the two together constituting the circuital relations of the electromagnetic field.

It has proved possible to obtain a solution of a very general type, by assuming that

$$(1\cdot1) \qquad X = \frac{\partial P}{\partial x} - xQ, \qquad Y = \frac{\partial P}{\partial y} - yQ, \qquad Z = \frac{\partial P}{\partial z} - zQ;$$

then equations (M) yield

$$(1\cdot2) \qquad -\mu\frac{\partial\alpha}{\partial t} = y\frac{\partial Q}{\partial z} - z\frac{\partial Q}{\partial y}, \qquad -\mu\frac{\partial\beta}{\partial t} = z\frac{\partial Q}{\partial x} - x\frac{\partial Q}{\partial z}, \qquad -\mu\frac{\partial\gamma}{\partial t} = x\frac{\partial Q}{\partial y} - y\frac{\partial Q}{\partial x}.$$

Substitute from equations (1·2) in the first equation (E) and we obtain

$$(1\cdot3) \qquad \frac{\mu K}{c^2}\frac{\partial^2 X}{\partial t^2} = \frac{\partial}{\partial x}\left(Q + x\frac{\partial Q}{\partial x} + y\frac{\partial Q}{\partial y} + z\frac{\partial Q}{\partial z}\right) - x\Delta^2 Q,$$

where Δ^2 denotes LAPLACE'S operator $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$.

On comparing equations (1·1) and (1·3) they will be seen to be consistent provided that

$$(1\cdot4) \qquad \frac{\mu K}{c^2}\frac{\partial^2 P}{\partial t^2} = Q + x\frac{\partial Q}{\partial x} + y\frac{\partial Q}{\partial y} + z\frac{\partial Q}{\partial z}$$

and that

$$(1\cdot5) \qquad \frac{\mu K}{c^2}\frac{\partial^2 Q}{\partial t^2} = \Delta^2 Q.$$

Thus Q must satisfy the fundamental wave-equation, which is satisfied by any component of the electric or magnetic forces (X, Y, Z) or (α, β, γ).

For our purpose it is more convenient to express the above solutions in terms of spherical polar co-ordinates r, θ, ϕ; these are supposed to form a right-handed system, when taken in this order, so as to avoid changes of sign in introducing the new co-ordinates. We write here* (R_1, R_2, R_3) for the components of electric force in the directions of r, θ, ϕ respectively; and (H_1, H_2, H_3) for the components of magnetic force.

Equations (1·1) then become

$$(1\cdot11) \qquad R_1 = \frac{\partial P}{\partial r} - rQ, \qquad R_2 = \frac{1}{r}\frac{\partial P}{\partial\theta}, \qquad R_3 = \frac{1}{r\sin\theta}\frac{\partial P}{\partial\phi}.$$

---

* This is done to avoid confusion with the Cartesian components used in equations (E) and (M); but in the subsequent sections we shall use (X, Y, Z) and (α, β, γ) for the spherical polar components here denoted by (R_1, R_2, R_3) and (H_1, H_2, H_3) respectively.

As regards the transformation of (1·1) to (1·11), it is sufficient to note that the gradient of P has the spherical polar components

$$\frac{\partial P}{\partial r}, \ \frac{1}{r}\frac{\partial P}{\partial \theta}, \ \frac{1}{r\sin\theta}\frac{\partial P}{\partial \phi},$$

and that $(r, 0, 0)$ corresponds to the Cartesian vector (x, y, z).

To obtain the formulæ corresponding to (1·2) we observe that the vector on the right is equal to the vector-product of the two vectors,

$$(x, y, z) \quad \text{and} \quad \left(\frac{\partial Q}{\partial x}, \frac{\partial Q}{\partial y}, \frac{\partial Q}{\partial z}\right);$$

and that these two are represented by

$$(r, 0, 0) \quad \text{and} \quad \left(\frac{\partial Q}{\partial r}, \frac{1}{r}\frac{\partial Q}{\partial \theta}, \frac{1}{r\sin\theta}\frac{\partial Q}{\partial \phi}\right).$$

Thus the vector-product has the spherical polar components

$$\left(0, \quad -\frac{1}{\sin\theta}\frac{\partial Q}{\partial \phi}, \quad \frac{\partial Q}{\partial \theta}\right).$$

Consequently equations (1·2) now become

$$(1·21) \qquad -\mu\frac{\partial H_1}{\partial t} = 0, \qquad -\mu\frac{\partial H_2}{\partial t} = -\frac{1}{\sin\theta}\frac{\partial Q}{\partial \phi}, \qquad -\mu\frac{\partial H_3}{\partial t} = \frac{\partial Q}{\partial \theta},$$

while (1·4) and (1·5) give

$$(1·41) \qquad \frac{\mu K}{c^2}\frac{\partial^2 P}{\partial t^2} = Q + r\frac{\partial Q}{\partial r} = \frac{\partial}{\partial r}(rQ).$$

$$(1·51) \qquad \frac{\mu K}{c^2}\frac{\partial^2 Q}{\partial t^2} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial Q}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial Q}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 Q}{\partial \phi^2}$$

$$= \frac{1}{r}\frac{\partial^2}{\partial r^2}(rQ) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial Q}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 Q}{\partial \phi^2}.$$

A consideration of these formulæ suggests that further simplifications can be obtained by writing

$$(1·6) \qquad P = \frac{\partial U}{\partial r}, \qquad rQ = \frac{\mu K}{c^2}\frac{\partial^2 U}{\partial t^2},$$

which together satisfy equation (1·41); and then equation (1·51) leads to the equation for U :—

$$(1·7) \qquad \frac{\mu K}{c^2}\frac{\partial^2 U}{\partial t^2} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta\frac{\partial U}{\partial \theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 U}{\partial \phi^2}.$$

Substituting from (1·6) in (1·11) and (1·21) we obtain the final expression for the field in terms of U:—

$$(1\cdot8)\quad \left\{\begin{aligned} R_1 &= \frac{\partial^2 U}{\partial r^2} - \frac{\mu K}{c^2}\frac{\partial^2 U}{\partial t^2} \\ R_2 &= \frac{1}{r}\frac{\partial^2 U}{\partial\theta\,\partial r} \\ R_3 &= \frac{1}{\rho}\frac{\partial^2 U}{\partial\phi\,\partial r} \end{aligned}\right. \qquad \begin{aligned} cH_1 &= 0 \\ cH_2 &= +\frac{1}{\rho}\frac{\partial}{\partial\phi}\left(\frac{K}{c}\frac{\partial U}{\partial t}\right), \\ cH_3 &= -\frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{K}{c}\frac{\partial U}{\partial t}\right). \end{aligned} \qquad \rho = r\sin\theta,$$

In like manner we obtain another set of solutions by making an assumption similar to (1·1) for the components of magnetic force (α, β, γ). This gives the field:—

$$(1\cdot9)\quad \left\{\begin{aligned} R_1 &= 0 \\ R_2 &= -\frac{1}{\rho}\frac{\partial}{\partial\phi}\left(\frac{\mu}{c}\frac{\partial V}{\partial t}\right) \\ R_3 &= +\frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{\mu}{c}\frac{\partial V}{\partial t}\right) \end{aligned}\right. \qquad \begin{aligned} cH_1 &= \frac{\partial^2 V}{\partial r^2} - \frac{\mu K}{c^2}\frac{\partial^2 V}{\partial t^2} \\ cH_2 &= \frac{1}{r}\frac{\partial^2 V}{\partial\theta\,\partial r}, \\ cH_3 &= \frac{1}{\rho}\frac{\partial^2 V}{\partial\phi\,\partial r}, \end{aligned} \qquad \rho = r\sin\theta,$$

where V is a second solution of equation (1·7).

It can be proved* that (1·8) gives the most general field in which the radial magnetic force (H_1) is zero, while (1·9) gives the most general field in which the radial electric force (R_1) is zero. It can also be shown that the field is uniquely determined by the value of R_1 and H_1; and accordingly the most general solution can be obtained by the superposition of (1·8) and (1·9).

§ 2. FURTHER SPECIALIZATION OF THE SOLUTION OF § 1.

If we superpose the fields (1·8), (1·9), and now utilize (X, Y, Z), (α, β, γ) to denote the spherical polar components of the field, we have the general solution† :—

$$(2\cdot1)\quad \left\{\begin{aligned} X &= \frac{\partial^2 U}{\partial r^2} - \frac{\mu K}{c^2}\frac{\partial^2 U}{\partial t^2} \\ Y &= \frac{1}{r}\frac{\partial^2 U}{\partial\theta\,\partial r} - \frac{1}{\rho}\frac{\partial}{\partial\phi}\left(\frac{\mu}{c}\frac{\partial V}{\partial t}\right), \\ Z &= \frac{1}{\rho}\frac{\partial^2 U}{\partial\phi\,\partial r} + \frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{\mu}{c}\frac{\partial V}{\partial t}\right), \end{aligned}\right. \qquad \rho = r\sin\theta,$$

---

* See a paper in the 'Philosophical Magazine,' July, 1919 (6th ser., vol. 38), p. 143.

† Originally worked out in 1899, and first published as a question in Part II. of the 'Mathematical Tripos,' 1910.

$$(2\cdot2)\qquad \left\{\begin{aligned} c\alpha &= \frac{\partial^2 V}{\partial r^2} - \frac{\mu K}{c^2}\frac{\partial^2 V}{\partial t^2} \\ c\beta &= \frac{1}{r}\frac{\partial^2 V}{\partial\theta\,\partial r} + \frac{1}{\rho}\frac{\partial}{\partial\phi}\left(\frac{K}{c}\frac{\partial U}{\partial t}\right), \qquad \rho = r\sin\theta, \\ c\gamma &= \frac{1}{\rho}\frac{\partial^2 V}{\partial\phi\,\partial r} - \frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{K}{c}\frac{\partial U}{\partial t}\right), \end{aligned}\right.$$

where U, V are any two solutions of the equation

$$(2\cdot3)\qquad \frac{\mu K}{c^2}\frac{\partial^2 U}{\partial t^2} = \frac{\partial^2 U}{\partial r^2} + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial U}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 U}{\partial\phi^2}.$$

A solution of (2·3) which is sufficiently general for the applications in view may be found by assuming that U and V can be expressed as sums of terms of the type

$$F(r, t)\times Y(\theta, \phi).$$

It is easy to see that then (2·3) leads to the equation

$$(2\cdot4)\qquad \frac{r^2}{F}\left(\frac{\partial^2 F}{\partial r^2} - \frac{\mu K}{c^2}\frac{\partial^2 F}{\partial t^2}\right) = -\frac{1}{Y}\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2}\right\},$$

and since the two sides of equation (2·4) are functions of r, t and of θ, ϕ respectively, it is clear that each side must be a mere constant. If we write the constant in the form $n(n+1)$, it is evident that Y must be a surface-harmonic of order n.

Accordingly in problems (such as those with which we shall be concerned in the sequel) where the *whole* of angular space is considered, the value of n must be a positive integer; for (except when n is an integer) there are no surface-harmonics which are everywhere continuous and single-valued.

Thus we may reduce our solution to the form

$$(2\cdot5)\qquad U \text{ or } V = \Sigma\, F_n(r, t)\, Y_n(\theta, \phi), \qquad n = 0, 1, 2, 3, \ldots,$$

where F_n is a solution of the equation

$$(2\cdot6)\qquad \frac{\partial^2 F_n}{\partial r^2} - \frac{1}{c_1^2}\frac{\partial^2 F_n}{\partial t^2} - \frac{n(n+1)}{r^2}F_n = 0,$$

and

$$c_1^2 = c^2/(\mu K).$$

The general solution of equation (2·6) is well known, and it is given by*

$$(2·7) \qquad F_n(r, t) = r^{n+1}\left(-\frac{1}{r}\frac{\partial}{\partial r}\right)^n \left\{\frac{f(c_1t-r)+g(c_1t+r)}{r}\right\},$$

where the functions f and g are arbitrary.

In the special case of *divergent waves*, the function g can be omitted in (2·7); and, if the region considered includes the origin, then $g(c_1t+r) = -f(c_1t+r)$, so as to make $F_n(r, t)$ continuous at $r = 0$.

It will be convenient to notice that in consequence of equation (2·6) the radial components of force can be written in the simpler forms

$$(2·8) \qquad X \quad \text{or} \quad c\alpha = \Sigma \frac{n(n+1)}{r^2} F_n(r, t) Y_n(\theta, \phi).$$

It will be noticed that we can at once determine the form (2·5) for U or V when the radial forces have been expressed in the form (2·8); this agrees with the general conclusion stated at the end of § 1, that (in spherical polar co-ordinates) the remaining components of force are completely determined when the two radial components are known.

§ 3. SPECIAL CASE OF SIMPLE HARMONIC WAVES AND THE APPROPRIATE FUNCTIONS.

We assume in future that the waves are simple harmonic, of wave-length $2\pi/\kappa$ in free space; we can then suppose the time to occur only in the form of a time-factor $e^{\iota\kappa ct}$, with the usual convention that finally only the real (or the imaginary) parts of the formulæ will be used.

The functions f, g occurring in equation (2·7) above are then exponentials of the types

$$e^{\iota\kappa_1(c_1t-r)} \quad \text{and} \quad e^{\iota\kappa_1(c_1t+r)},$$

where κ_1 is given by

$$\kappa_1 c_1 = \kappa c, \quad \text{or} \quad \kappa_1 = \kappa\sqrt{(\mu K)}.$$

Thus (if we now suppress the time-factor $e^{\iota\kappa ct}$) the functions given by (2·7) are of the types

$$(3·1) \qquad r^{n+1}\left(-\frac{1}{r}\frac{\partial}{\partial r}\right)^n \frac{e^{-\iota\kappa_1 r}}{r}, \qquad r^{n+1}\left(-\frac{1}{r}\frac{\partial}{\partial r}\right)^n \frac{e^{+\iota\kappa_1 r}}{r}.$$

We shall be concerned with two special types only: (i.) divergent waves; (ii.) waves which are continuous at $r = 0$. The former of these corresponds to the

---

[*] See, for instance, LAMB'S 'Hydrodynamics,' 1906, art. 295; an alternative method of solution given in § 3 of my paper in the 'Philosophical Magazine,' quoted on p. 179 above; compare also A. E. H. LOVE ('Phil. Trans. Roy. Soc.,' A, vol. 197, 1901, pp. 9, 10).

first expression in (3·1); while the latter is found by combining the two expressions so as to yield

$$(3{\cdot}2) \qquad r^{n+1}\left(-\frac{1}{r}\frac{\partial}{\partial r}\right)^n \frac{\sin(\kappa_1 r)}{r}.$$

Using the notation explained in (3·4) and (3·5) below, the standard functions are

$E_n(\kappa_1 r)$ for divergent waves,

and

$S_n(\kappa_1 r)$ for waves within a spherical boundary.

Consequently, for waves inside a spherical boundary, (2·5) and (2·8) can now be replaced by the forms

$$(3{\cdot}3) \qquad \begin{cases} \mathrm{U} \quad \text{or} \quad \mathrm{V} = \Sigma S_n(\kappa_1 r)\, Y_n(\theta, \phi), \\ \mathrm{X} \quad \text{or} \quad c\alpha = \Sigma \dfrac{n(n+1)}{r^2} S_n(\kappa_1 r)\, Y_n(\theta, \phi), \end{cases}$$

for divergent waves the function $S_n(\kappa_1 r)$ must be replaced by $E_n(\kappa_1 r)$.

Definitions and Properties of the Two Standard Functions $S_n(z)$, $E_n(z)$.

We write for brevity

$$(3{\cdot}4) \qquad S_n(z) = z^{n+1}\left(-\frac{1}{z}\frac{d}{dz}\right)^n\left(\frac{\sin z}{z}\right)$$

$$= \frac{z^{n+1}}{1\,.\,3\,.\,5\ldots(2n+1)}\left\{1 - \frac{z^2}{2(2n+3)} + \frac{z^4}{2\,.\,4(2n+3)(2n+5)} - \ldots\right\}.$$

In terms of the known Bessel function we can write

$$(3{\cdot}41) \qquad S_n(z) = \sqrt{\left(\frac{\pi z}{2}\right)}\, J_{n+\frac{1}{2}}(z),$$

and accordingly the function $S_n(z)$ is the same as that denoted by u in one of MACDONALD's papers.*

In the notation adopted by LAMB,† and those writers who have used LAMB's solutions as the fundamental forms, we have the identity

$$(3{\cdot}42) \qquad S_n(z) = z^{n+1}\psi_n(z).$$

---

* 'Phil. Trans. Roy. Soc.,' A, vol. 210, 1910, p. 113. See in particular p. 115.

† 'Hydrodynamics,' 1906, Art. 287.

Similarly, we write

(3·5)
$$\mathrm{E}_n(z) = z^{n+1}\left(-\frac{1}{z}\frac{d}{dz}\right)^n\left(\frac{e^{-\iota z}}{z}\right)$$
$$= \mathrm{C}_n(z) - \iota\mathrm{S}_n(z),$$

where

(3·6)
$$\mathrm{C}_n(z) = z^{n+1}\left(-\frac{1}{z}\frac{d}{dz}\right)^n\left(\frac{\cos z}{z}\right)$$
$$= \frac{1\,.\,3\ldots(2n-1)}{z^n}\left\{1 - \frac{z^2}{2(1-2n)} + \frac{z^4}{2\,.\,4(1-2n)(3-2n)} - \ldots\right\}.$$

In terms of the K_n function (the modified Bessel function used by MACDONALD) we have the relation

(3·51)
$$\mathrm{E}_n(z) = \sqrt{\left(\frac{2z}{\pi}\right)}\, e^{\frac{1}{2}(n+\frac{1}{2})\pi\iota}\mathrm{K}_{n+\frac{1}{2}}(\iota z).$$

Thus

$$\mathrm{E}_n(z) = v - \iota u$$

in terms of the notation used by MACDONALD in the paper last quoted, and in LAMB's notation

(3·7)
$$\mathrm{E}_n(z) = z^{n+1}\{\Psi_n(z) - \iota\psi_n(z)\} = z^{n+1}f_n(z).$$

In consequence of the equation (2·6) we see that both $\mathrm{S}_n(z)$ and $\mathrm{E}_n(z)$ are solutions of the differential equation

(3·8)
$$\frac{d^2\mathrm{S}_n}{dz^2} + \left\{1 - \frac{n(n+1)}{z^2}\right\}\mathrm{S}_n = 0,$$

The functions $\mathrm{S}_n(z)$, $\mathrm{C}_n(z)$ and $|\mathrm{E}_n(z)|$ have been tabulated from $z = 1$ to 10, and for values of n ranging from 0 to 22, by Mr. DOODSON,* and these tables have formed the basis of the numerical calculations mentioned on p. 176 above.†

It will be convenient to collect here the simple relations amongst the functions S_{n-1}, S_n, S_{n+1}, which correspond to the known results for Bessel functions, or to those given by LAMB for the equivalent function $\psi_n(z)$.

Difference Relations for the Functions S_n, E_n.

From (3·4) we see that

(3·81)
$$\mathrm{S}_{n+1}(z) = -z^{n+1}\frac{d}{dz}\left\{\frac{\mathrm{S}_n(z)}{z^{n+1}}\right\} = \frac{n+1}{z}\mathrm{S}_n(z) - \frac{d\mathrm{S}_n}{dz},$$

and by using (3·5) we see that the same relation holds for $\mathrm{E}_n(z)$.

Again, it will be found that

$$\left(\frac{d}{dz} + \frac{n}{z}\right)z^{n+p} = (2n+p)\,z^{n+p-1},$$

---

* 'British Association Report,' 1914.

† PROUDMAN, DOODSON and KENNEDY, 'Phil. Trans. Roy. Soc.,' A, vol. 217, 1917, p. 279.

and using this in equation (3·4) we deduce that

$$(3\cdot82) \qquad S_{n-1}(z) = \left(\frac{d}{dz} + \frac{n}{z}\right) S_n(z) = \frac{n}{z} S_n(z) + \frac{dS_n}{dz}.$$

Combining (3·81) and (3·82) we have also

$$(3\cdot83) \qquad S_{n-1}(z) + S_{n+1}(z) = \frac{2n+1}{z} S_n(z).$$

The relations (3·82), (3·83) hold equally for $E_n(z)$ and $C_n(z)$, as may be seen from (3·5) and (3·6). As $E_n(z)$, $S_n(z)$ are independent solutions of the equation (3·8), it is evident that

$$E_n(z) \frac{dS_n}{dz} - S_n(z) \frac{dE_n}{dz} = \text{const.}$$

Now when z is small, it is easy to verify from (3·4) and (3·6) that

$$E_n(z) \frac{dS_n}{dz} = \frac{n+1}{2n+1} + O(z^2), \qquad S_n(z) \frac{dE_n}{dz} = -\frac{n}{2n+1} + O(z^2),$$

and accordingly we have

$$(3\cdot84) \qquad E_n(z) \frac{dS_n}{dz} - S_n(z) \frac{dE_n}{dz} = 1.$$

In the discussions of § 6, when n, z are both large, it will be convenient to adopt the following notation :—

$$(3\cdot85) \qquad |E_n(z)| = R, \quad E_n(z) = Re^{-\iota\psi}, \quad \text{so that} \quad S_n(z) = R \sin \psi, \quad C_n(z) = R \cos \psi.$$

Substituting from (3·85) in (3·84) we deduce that

$$(3\cdot86) \qquad R^2 \frac{d\psi}{dz} = 1.$$

Before leaving these preliminary formulæ it will be convenient to quote the formula for $e^{\iota\kappa z}$ in terms of our standard functions; namely

$$(3\cdot9) \qquad e^{\iota\kappa z} = \sum_{n=0}^{\infty} (2n+1) \iota^n \frac{S_n(\kappa r)}{\kappa r} P_n(\mu),$$

where $z = r \cos \theta = r\mu$ and $P_n(\mu)$ is LEGENDRE'S polynomial of order n.

This result follows at once from the formula given in LAMB'S 'Hydrodynamics,' Art. 291, on using the relation (3·4) between $\psi_n(\kappa r)$ and $S_n(\kappa r)$, already quoted. It is of course evident that an expansion of the type (3·9) might be anticipated, since each side satisfies the wave-equation, is symmetrical about the axis of z, and is continuous at $r = 0$; the determination of the numerical coefficients may be then carried out quickly by comparing the terms in $(\kappa r\mu)^n$ on the two sides of the equation.

§ 4. PLANE ELECTROMAGNETIC WAVES INCIDENT ON A SPHERICAL OBSTACLE.

Suppose that the incident wave-train is travelling along the negative direction of the axis of z (that is, from $\theta = 0$ towards $\theta = \pi$); and that it is polarized in the plane of yz (that is, in the plane $\phi = \frac{1}{2}\pi$). Suppose further that the electric force in the wave-train has unit amplitude; then, in terms of the Cartesian specification, the incident wave is defined by*

$$X = -e^{\iota\kappa(ct+z)}, \qquad c\beta = +e^{\iota\kappa(ct+z)},$$

the remaining components of force being zero.

We must first express this wave in the standard forms of (2·1) and (2·2); we therefore introduce polar co-ordinates, and then proceed to find the radial components of force, which will suffice to determine the functions U, V.

These radial components are given by

$$(4·1) \qquad X = -\sin\theta\cos\phi\, e^{\iota\kappa r\cos\theta}, \qquad c\alpha = +\sin\theta\sin\phi\, e^{\iota\kappa r\cos\theta},$$

where the time-factor $e^{\iota\kappa ct}$ is now omitted.

Now from (3·9) we have the formula

$$e^{\iota\kappa r\cos\theta} = \sum_{n=0}^{\infty}(2n+1)\,\iota^{n}\,\frac{S_n(\kappa r)}{\kappa r}\,P_n(\cos\theta).$$

So, differentiating with regard to θ, we find that

$$(4·2) \qquad \sin\theta e^{\iota\kappa r\cos\theta} = -\sum_{n=1}^{\infty}(2n+1)\,\iota^{n-1}\,\frac{S_n(\kappa r)}{(\kappa r)^2}\,P'_n(\cos\theta)\sin\theta.$$

Accordingly, on substituting (4·2) in (4·1), we find that in the incident wave

$$(4·3) \quad \left\{ \begin{aligned} U &= -\frac{\sin\theta\cos\phi}{\kappa^2}\sum_{n=1}^{\infty}\frac{(2n+1)}{n(n+1)}\,\iota^{n-1}S_n(\kappa r)\,P'_n(\cos\theta), \\ &\text{and} \\ V &= +\frac{\sin\theta\sin\phi}{\kappa^2}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\,\iota^{n-1}S_n(\kappa r)\,P'_n(\cos\theta) \end{aligned} \right.$$

by comparing the two formulæ (2·5) and (2·8).

The corresponding waves in the interior of the sphere will be given by the two functions

$$(4·4) \quad \left\{ \begin{aligned} U_1 &= -\frac{\sin\theta\cos\phi}{\kappa^2}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\,\iota^{n-1}B_nS_n(\kappa_1 r)\,P'_n(\cos\theta), \\ &\text{and} \\ V_1 &= +\frac{\sin\theta\cos\phi}{\kappa^2}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\,\iota^{n-1}D_nS_n(\kappa_1 r)\,P'_n(\cos\theta), \end{aligned} \right.$$

where

$$\kappa_1 = \kappa\sqrt{(\mu K)}$$

and K, μ are the fundamental constants of the spherical obstacle.

---

* It is assumed that in the incident wave we may take $\mu = 1$, $K = 1$, $c_1 = c$, $\kappa_1 = \kappa$.

Similarly the scattered waves will be given by the two functions

$$
(4{\cdot}5)\quad \left\{\begin{aligned}
U_0 &= -\frac{\sin\theta\cos\phi}{\kappa^2}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\,\iota^{n-1}A_nE_n(\kappa r)\,P'_n(\cos\theta)\\
&\text{and}\\
V_0 &= +\frac{\sin\theta\sin\phi}{\kappa^2}\sum_{n=1}^{\infty}\frac{2n+1}{n(n+1)}\,\iota^{n-1}C_nE_n(\kappa r)\,P'_n(\cos\theta).
\end{aligned}\right.
$$

The boundary conditions are given by the continuity of the tangential components of electric and magnetic force at the sphere $r = a$.

It is evident from the form of equations (2·1), (2·2) that these conditions will be satisfied if we take

$$
(4{\cdot}6)\quad \left\{\begin{aligned}
&\frac{\partial U_1}{\partial r} = \frac{\partial U}{\partial r}+\frac{\partial U_0}{\partial r}, \qquad KU_1 = U+U_0\\
&\text{and}\\
&\frac{\partial V_1}{\partial r} = \frac{\partial V}{\partial r}+\frac{\partial V_0}{\partial r}, \qquad \mu V_1 = V+V_0.
\end{aligned}\right.
$$

Thus we find that A_n and C_n (the coefficients in the scattered waves) are given by

$$
(4{\cdot}7)\quad \left\{\begin{aligned}
&\{S_n(\kappa a)+A_nE_n(\kappa a)\}\,\frac{\kappa_1}{\kappa K}\,\frac{S'_n(\kappa_1 a)}{S_n(\kappa_1 a)} = S'_n(\kappa a)+A_nE'_n(\kappa a)\\
&\text{and}\\
&\{S_n(\kappa a)+C_nE_n(\kappa a)\}\,\frac{\kappa_1}{\kappa\mu}\,\frac{S'_n(\kappa_1 a)}{S_n(\kappa_1 a)} = S'_n(\kappa a)+C_nE'_n(\kappa a).
\end{aligned}\right.
$$

The special case of a *perfectly conducting sphere* is given by making the tangential electric force zero at the sphere $r = a$; and this condition is satisfied if

$$
(4{\cdot}61)\qquad 0 = \frac{\partial U}{\partial r}+\frac{\partial U_0}{\partial r}, \qquad 0 = V+V_0.
$$

Thus we find the simpler formulæ

$$
(4{\cdot}71)\qquad S'_n(\kappa a)+A_nE'_n(\kappa a) = 0 \qquad S_n(\kappa a)+C_nE_n(\kappa a) = 0,
$$

which may be regarded as limiting forms of (4·7), when $|K| \to \infty$ and $\mu \to 0$.

The formulæ (4·5) and (4·71) lead at once to those quoted by Dr. J. PROUDMAN* in calculating the pressure of radiation due to a plane wave incident on a small conducting sphere.

In all the applications with which we shall be concerned at present the point at which the disturbance is to be calculated will be at a distance large compared with the wave-length. Then we can simplify the general formulæ by observing that (3·5) may be replaced by the approximation

---

* 'Monthly Notices of the Royal Astronomical Society,' vol. 73, 1913, p. 535.

$$E_n(\kappa r) = \iota^n e^{-\iota\kappa r}$$

if $1/\kappa r$, $(1/\kappa r)^2$, &c., are neglected. Further, in the final formulæ for the forces, U_0 and V_0 occur only in the two combinations

$$M = \frac{1}{r}\frac{\partial U_0}{\partial r} = -\frac{\iota\kappa U_0}{r}, \qquad N = -\frac{1}{r}\frac{\partial V_0}{\partial r} = \frac{\iota\kappa V_0}{r},$$

where terms of the relative order $1/\kappa r$ have been rejected.

On substituting from (4·5) we find, to the same degree of accuracy,

$$(4\cdot8)\quad \left\{\begin{aligned} M &= +\frac{1}{r}\frac{\partial U_0}{\partial r} = -\frac{\iota\kappa}{r}U_0 = \sin\theta\cos\phi\,\frac{e^{-\iota\kappa r}}{\kappa r}\sum_{n=1}^{\infty}(-1)^n\frac{2n+1}{n(n+1)}A_nP'_n(\cos\theta) \\ &\text{and} \\ N &= -\frac{1}{r}\frac{\partial V_0}{\partial r} = +\frac{\iota\kappa}{r}V_0 = \sin\theta\sin\phi\,\frac{e^{-\iota\kappa r}}{\kappa r}\sum_{n=1}^{\infty}(-1)^n\frac{2n+1}{n(n+1)}C_nP'_n(\cos\theta). \end{aligned}\right.$$

Then, substituting in the general formulæ (2·1) and (2·2), we find that (to our present order of approximation) the radial components of force are zero, and that the transverse components are given by

$$(4\cdot9)\quad \left\{\begin{aligned} Y &= \frac{\partial M}{\partial\theta} - \frac{1}{\sin\theta}\frac{\partial N}{\partial\phi} = +c\gamma \\ Z &= \frac{1}{\sin\theta}\frac{\partial M}{\partial\phi} + \frac{\partial N}{\partial\theta} = -c\beta. \end{aligned}\right.$$

Accordingly *the electric and magnetic forces in the scattered waves are at right angles to each other and to the radius, and their magnitudes are related in the same manner as in a plane wave.*

This conclusion might very well have been anticipated; and for the case of *small* obstacles of any shape (with constants K, μ differing but little from unity) the conclusion is contained in a paper by Lord RAYLEIGH.* But I cannot find that it has been noticed for the case of spheres of any size, and of any electrical and magnetic constants.

This may serve to indicate one advantage of the formulæ in spherical polars over those in Cartesian co-ordinates.

The formulæ (4·8), (4·9), with the values of A_n, C_n given by (4·71), were those used by Messrs. PROUDMAN, DOODSON, and KENNEDY in their numerical calculations quoted in the introduction to this paper.

---

* 'Scientific Papers,' vol. 1, pp. 522–536. For a *small* perfectly conducting sphere the same conclusion is given by Sir J. J. THOMSON, 'Recent Researches,' p. 448.

§ 5. SPHERES SMALL COMPARED WITH THE WAVE-LENGTH.

The fundamental assumption is that $|\kappa a|$ is small enough to justify us in rejecting all but one or two terms in the power-series for $S_n(\kappa a)$ and $E_n(\kappa a)$.

It has been usual to assume further that $|\kappa_1 a|$ is correspondingly small; but Dr. PROUDMAN has remarked that in the case of a dielectric sphere with a large value of K the second assumption need not follow from the first. It seems worth while therefore to simplify the formulæ (4·7) by expanding in powers of $|\kappa a|$, while retaining the general forms for $S_n(\kappa_1 a)$; it will be seen moreover that the resulting formulæ form a link between the results for dielectric spheres and those for conductors.

Remembering that $S_n(\kappa a)$ is of order $(\kappa a)^{n+1}$ and that $E_n(\kappa a)$ is of order $(\kappa a)^{-n}$, it is easy to see from (4·7) that in general A_n and C_n are both of order $(\kappa a)^{2n+1}$. Thus in the first approximation it will be sufficient to deal only with the coefficients A_1 and C_1; and for these we need the formulæ for $S_1(\kappa_1 a)$ and $S'_1(\kappa_1 a)$. Now from (3·4) we have

$$S_1(z) = z^2\left(-\frac{1}{z}\frac{d}{dz}\right)\left(\frac{\sin z}{z}\right) = \frac{\sin z}{z} - \cos z = \frac{\sin z}{z}(1 - z\cot z),$$

and so

$$S'_1(z) = \sin z\left(1 - \frac{1}{z^2}\right) + \frac{\cos z}{z} = \frac{\sin z}{z^2}(z^2 - 1 + z\cot z).$$

Hence

$$\frac{zS'_1(z)}{S_1(z)} = \frac{z^2 - 1 + z\cot z}{1 - z\cot z} = F(z) - 1,$$

where now

$$F(z) = z^2/(1 - z\cot z).$$

In the second place, for the functions of κa, from (3·4) and (3·6) we find the first approximations

$$S_1(z) = \tfrac{1}{3}z^3, \qquad E_1(z) = 1/z.$$

Substituting, it will be seen that for $n = 1$ the first equation in (4·7) gives

$$\left\{\frac{3A_1}{(\kappa a)^3} + 1\right\}\frac{q}{K} = 2 - \frac{3A_1}{(\kappa a)^3},$$

where

$$q = \frac{\kappa_1 a S'_1(\kappa_1 a)}{S_1(\kappa_1 a)} = F(\kappa_1 a) - 1.$$

After a little reduction the last equation gives

$$(5·1) \qquad A_1 = \tfrac{1}{3}(\kappa a)^3 \frac{2K + 1 - F(\kappa_1 a)}{K - 1 + F(\kappa_1 a)}.$$

The second equation in (4·7) gives a similar formula for C_1, with μ taking the place of K. However, in most cases, it is sufficient to write $\mu = 1$, and then the formula simplifies further and becomes

$$(5{\cdot}2) \qquad C_1 = (\kappa a)^3 \left\{ \frac{1}{F(\kappa_1 a)} - \tfrac{1}{3} \right\}.$$

The two formulæ (5·1) and (5·2) are due to Dr. PROUDMAN, who has pointed out that they connect the results found by Lord RAYLEIGH for the case of dielectric spheres, and by Sir J. J. THOMSON for conducting spheres.

To deal with the case of dielectric spheres we do not regard K as large, so that $\kappa_1 a$ may be regarded as small (of the same order as κa); and then the approximations

$$1 - (\kappa_1 a) \cot (\kappa_1 a) = \tfrac{1}{3} (\kappa_1 a)^2, \qquad \text{or} \qquad F(\kappa_1 a) = 3,$$

may be used. This gives, in place of (5·1), (5·2) the simpler forms due to Lord RAYLEIGH*

$$(5{\cdot}21) \qquad A_1 = \tfrac{2}{3} (\kappa a)^3 \frac{K-1}{K+2}, \qquad C_1 = 0.$$

On the other hand, Sir J. J. THOMSON'S case corresponds to the assumption that K is of the form $K_1 - \iota K_2$ where K_2 is very large; then $|\kappa_1 a|$ may be regarded as large, and $\kappa_1 a$ as complex, with a negative imaginary part. Thus approximately† $\cot (\kappa_1 a) = \iota$, and so $|F(\kappa_1 a)| = |\kappa_1 a|$, which (although large) is small compared with $K| = |\kappa_1 a|^2/(\kappa a)^2$. Hence we find from (5·1) and (5·2) the approximate results,

$$(5{\cdot}22) \qquad A_1 = \tfrac{2}{3} (\kappa a)^3, \qquad C_1 = \tfrac{1}{3} (\kappa a)^3$$

as given by Sir J. J. THOMSON.‡ Of course this pair of formulæ follow at once from (4·71), on inserting the approximations for $S_1(\kappa a)$ and $E_1(\kappa a)$ given on p. 188 above.

Dr. PROUDMAN makes the further remark that, under the conditions assumed in (5·1) and (5·2), variations in the wave-length may produce very considerable changes in the magnitudes of A_1 and C_1, on account of the presence in $F(\kappa_1 a)$ of $\cot(\kappa_1 a)$, which may vary very fast. It is of course supposed that the sphere is dielectric, otherwise $\cot(\kappa_1 a)$ could be replaced by ι, as already stated.†

It is worth while to note the simple formulæ for the scattered wave, derived from (4·9); these give, to the present order of approximation

$$(5{\cdot}3) \qquad \begin{cases} Y = +c\gamma = \dfrac{e^{-\iota\kappa r}}{\kappa r} \left(\tfrac{3}{2} C_1 - \tfrac{3}{2} A_1 \cos \theta\right) \cos \phi, \\[2ex] Z = -c\beta = \dfrac{e^{\iota\kappa r}}{\kappa r} \left(\tfrac{3}{2} A_1 - \tfrac{3}{2} C_1 \cos \theta\right) \sin \phi ; \end{cases}$$

---

* 'Scientific Papers,' vol. 4, p. 321 (106); see also vol. 1, p. 526.

† Provided that the imaginary part of $\kappa_1 a$ exceeds π in numerical value, the error in this approximation is less than one half per cent.

‡ 'Recent Researches,' p. 148.

and so (to this order) the scattered wave is zero in the direction given by

$$(5{\cdot}31) \qquad \phi = 0, \qquad \cos\theta = C_1/A_1,$$

provided that A_1 is numerically greater than C_1. Thus in Lord RAYLEIGH'S case, the direction is given by $\theta = \frac{1}{2}\pi$;* and in Sir J. J. THOMSON'S by $\theta = \frac{2}{3}\pi$.

It may be noted here that, if the sphere has a sufficiently large dielectric constant K, it may happen that A_1 is numerically less than C_1; and then the direction given by (5·31) is no longer real.

Taking K to be real (the case of a conductor having been already considered on p. 189), it is easy to see that $A_1 < C_1$ gives $F(\kappa_1 a) < 1$ (on the assumption $K > 1$). Now the function $F(z)$ steadily decreases from 3 to 0 as z varies from 0 to π; and a rough calculation shows that for $z = \frac{7}{8}\pi$, $F(z)$ is slightly less than unity. Also in order to justify the approximations used for $S_1(\kappa a)$ and $E_1(\kappa a)$, we must suppose that $(\kappa a)^2 \leqq \frac{1}{10}$.

Hence the possibility contemplated may occur if, say,

$$(\tfrac{7}{8}\pi)^2 \leqq (\kappa_1 a)^2 \leqq \pi^2, \qquad \text{and} \qquad (\kappa a)^2 \leqq \tfrac{1}{10},$$

giving

$$K \geqq 75.$$

The direction in which the scattered wave vanishes will be given by

$$(5{\cdot}32) \qquad \phi = \tfrac{1}{2}\pi, \qquad \cos\theta = A_1/C_1 = 2F(\kappa_1 a)/\{3 - F(\kappa_1 a)\},$$

the final formula being simplified by remembering that K is large.

I am not aware that there is any experimental evidence showing traces of this phenomenon; in fact all the evidence shows that $\phi = 0$, $\theta = \frac{1}{2}\pi$ is not far from the truth. Thus the circumstances in actual experiments cannot have been such as to introduce the reversal of magnitude between A_1 and C_1.

(ii.) *Second Approximations.*

We proceed next to find second approximations, assuming that $|K|$ is not large; it will be necessary to retain the second terms in the series (3·4) and (3·6) for S_1 and E_1, but the first terms will suffice for S_2 and E_2.† It is easy to see that then terms of order $(\kappa a)^5$ occur in the coefficients A_1, C_1 and A_2, but that no other coefficients can contain terms of order lower than $(\kappa a)^7$.

Using now the series (3·4) for $S_1(z)$ we have

$$S_1(z) = \tfrac{1}{3}z^2(1 - \tfrac{1}{10}z^2), \qquad S'_1(z) = \tfrac{2}{3}z(1 - \tfrac{1}{5}z^2),$$

retaining the second terms only in each series. Thus we find, to the same order,

$$z\frac{S'_1(z)}{S_1(z)} = 2 - \tfrac{1}{5}z^2.$$

Also (3·6) gives similarly

$$E_1(z) = \frac{1}{z}(1 + \tfrac{1}{2}z^2), \qquad E'_1(z) = -\frac{1}{z^2}(1 - \tfrac{1}{2}z^2).$$

---

* A closer approximation is worked out on the next page; see formulæ (5·6) below.

† It would be possible, of course, to obtain second approximations to (5·1) and (5·2), but a glance at the formulæ shows that the work is so laborious as to be almost impracticable.

On substituting these results (4·7) becomes

$$\left(\frac{2}{\mathrm{K}}-\frac{p^2}{5}\right)\left\{\frac{p^2}{3}\left(1-\frac{p^2}{10}\right)+\frac{\mathrm{A}_1}{p}\left(1+\tfrac{1}{2}p^2\right)\right\}=\frac{2p^2}{3}\left(1-\frac{p^2}{5}\right)-\frac{\mathrm{A}_1}{p}\left(1-\frac{p^2}{2}\right)$$

and

$$\left(2-\frac{\mathrm{K}p^2}{5}\right)\left\{\frac{p^2}{3}\left(1-\frac{p^2}{10}\right)+\frac{\mathrm{C}_1}{p}\left(1+\tfrac{1}{2}p^2\right)\right\}=\frac{2p^2}{3}\left(1-\frac{p^2}{5}\right)-\frac{\mathrm{C}_1}{p}\left(1-\frac{p^2}{2}\right),$$

where, for brevity, we have written

$$\kappa a = p, \qquad (\kappa_1 a)^2 = \mathrm{K}p^2.$$

On reducing these equations, the results are

$$(5{\cdot}4)\quad \left\{\begin{aligned} &\mathrm{A}_1 = \tfrac{2}{3}p^3\left(\frac{\mathrm{K}-1}{\mathrm{K}+2}\right)\left\{1+\tfrac{3}{5}\left(\frac{\mathrm{K}-2}{\mathrm{K}+2}\right)p^2\right\} \\ &\text{and} \\ &\mathrm{C}_1 = \tfrac{1}{45}\,(\mathrm{K}-1)\,p^5. \end{aligned}\right.$$

To determine A_2, the series for $\mathrm{S}_2(z)$ and $\mathrm{E}_2(z)$ will be required to the first terms only; these are

$$\mathrm{S}_2(z)=\frac{z^3}{15}, \qquad \mathrm{E}_2(z)=\frac{3}{z^2},$$

and on substituting in (4·7), we find

$$(5{\cdot}5)\qquad \mathrm{A}_2=\frac{p^5}{15}\left(\frac{\mathrm{K}-1}{2\mathrm{K}+3}\right).$$

The field of the scattered waves is then given by

$$(5{\cdot}6)\quad \left\{\begin{aligned} \mathrm{Y} &= +c\gamma = \frac{e^{-\iota\kappa r}}{\kappa r}\left(\tfrac{3}{2}\mathrm{C}_1-\tfrac{3}{2}\mathrm{A}_1\cos\theta+\tfrac{5}{2}\mathrm{A}_2\cos 2\theta\right)\cos\phi, \\ \mathrm{Z} &= -c\beta = \frac{e^{-\iota\kappa r}}{\kappa r}\left(\tfrac{3}{2}\mathrm{A}_1-\tfrac{3}{2}\mathrm{C}_1\cos\theta-\tfrac{5}{2}\mathrm{A}_2\cos\theta\right)\sin\phi. \end{aligned}\right.$$

The field (5·6) *is accordingly zero* (*to the same degree of approximation*) *in the direction given by*

$$\phi = 0, \qquad \cos\theta = \left(\mathrm{C}_1-\tfrac{5}{3}\mathrm{A}_2\right)/\mathrm{A}_1 = \tfrac{1}{15}\,\frac{(\mathrm{K}-1)(\mathrm{K}+2)}{2\mathrm{K}+3}\,(\kappa a)^2.$$

This conclusion is apparently new; but it confirms an approximate result due to Lord RAYLEIGH,[*] according to which the scattered wave is zero in the direction given by

$$\phi = 0, \qquad \cos\theta = \tfrac{1}{25}\,(\mathrm{K}-1)\,(\kappa a)^2,$$

when (K−1) *is treated as small.* But on the other hand, our result contradicts a statement made by Prof. LOVE* that there is no direction in which the scattered wave is completely cut out; however, on a closer examination of Prof. LOVE's formulæ, they appear to confirm the present conclusion.

---

[*] 'Scientific Papers,' vol. 1, p. 531, formula (61).

The formulæ in question are (42) and (43) of the paper just quoted, but apparently there is a slip in (43). In the last line of (43) the factor given as 0 should really be $(z^2-y^2)/r^2$; the source of the inaccuracy being apparently in the passage from the formula (39) to (41). On introducing this additional term in the magnetic force, it appears that the electric and magnetic forces are zero in the direction given by†

$$x = 0, \qquad \frac{z}{r} = \frac{(K+2)(K-1)}{15(2K+3)}(\kappa a)^2$$

in Prof. LOVE'S notation; of course "zero" means that the forces are really of order $(\kappa a)^7$ at most.

It is not difficult to prove that the formulæ (5·6) agree with those found by Lord RAYLEIGH‡ and Prof. LOVE;§ the method to be adopted is similar to that used in § 4 of my paper in the 'Philosophical Magazine' (quoted on p. 179 above). But it should be observed that in the specification of the incident wave adopted by Lord RAYLEIGH and Prof. LOVE, the electric force is parallel to the axis of y; but here the electric force is parallel to the negative direction of x. Thus if ϕ' denotes the azimuthal angle corresponding to the former specification, it is evident that $\phi' = \frac{1}{2}\pi$ corresponds to $\phi = \pi$; and accordingly we shall have in general the relation

$$\phi - \phi' = \tfrac{1}{2}\pi,$$

because both angles are measured in the right-handed sense about the axis of z.

Lord RAYLEIGH'S paper contains tables and graphs from which it is easy to determine the variation of the field with θ; and in order to connect his tables with our formulæ, let us write (5·6) in the form

$$(5·7) \quad \left\{ \begin{aligned} & \mathrm{Y} = +c\gamma = \frac{e^{-\iota\kappa r}}{\kappa r}\mathrm{R}\cos\phi, \qquad \mathrm{Z} = -c\beta = \frac{e^{-\iota\kappa r}}{\kappa r}\mathrm{S}\sin\phi, \\ & \text{where} \\ & \mathrm{R} = \tfrac{3}{2}\mathrm{C}_1 - \tfrac{3}{2}\mathrm{A}_1\cos\theta + \tfrac{5}{2}\mathrm{A}_2\cos 2\theta, \\ & \mathrm{S} = \tfrac{3}{2}\mathrm{A}_1 - \tfrac{3}{2}\mathrm{C}_1\cos\theta - \tfrac{5}{2}\mathrm{A}_2\cos\theta. \end{aligned} \right.$$

Consider, first, points in the plane given by $x = 0$, in Lord RAYLEIGH'S notation; this gives $\phi' = \frac{1}{2}\pi$, or $\phi = \pi$. Hence, omitting the factor $e^{-\iota\kappa r}/(\kappa r)$, the electric force is equal to R (in the direction of θ *decreasing*); and accordingly R *is represented by the graph of the Cartesian component* $(y\mathrm{Z}-z\mathrm{Y})/r$, given by Lord RAYLEIGH.

Secondly, consider the plane $y = 0$; that is $\phi' = 0$, or $\phi = \frac{1}{2}\pi$. Here the electric force is equal to S, perpendicular to the plane; thus S *is represented by the graph of the Cartesian component* Y.

---

* 'Proc. Lond. Math. Soc.,' vol. 30, 1899, p. 318.

† A numerical slip in the first line of each of the formulæ (42) and (43) has to be corrected; the correction was given in the *Errata*, vol. 31, 'Proc. Lond. Math. Soc.'

‡ 'Scientific Papers,' vol. 5, p. 559, (u) and (v).

§ Formulæ (42) and (43) of the paper just quoted (allowing for the corrections just mentioned).

In general, the resultant electric force is represented by

$$\frac{e^{-\iota\kappa r}}{\kappa r}\sqrt{(R^2\cos^2\phi+S^2\sin^2\phi)}.$$

§ 5. (iii.) *Second Approximations for Conducting Spheres.*

The foregoing algebra needs no alteration beyond replacing $\sqrt{K}$ by the appropriate complex refractive index associated with the particular metal and wave-length considered. This of course assumes that $|K|$ is not so large that the convergence of $S_1(\kappa_1 a)$ becomes too slow to justify the approximation made above; and then the formulæ (5·4) to (5·6) provide the solution. It will be noticed that when K is complex, the equation

$$\cos\theta = \tfrac{1}{15}\frac{(K-1)(K+2)}{2K+3}(\kappa a)^2$$

will not usually give a real value for θ: and so *there is usually no direction in which the scattered wave is zero*[*] (or of order $(\kappa a)^7$).

It may be of interest to note here that experimental work on the scattering of light by fine particles has been carried out with silver particles suspended in water.[†] The corresponding values of κa seem to vary from $\frac{1}{2}$ to 2, and the value of $\sqrt{K}$ is taken as $0{\cdot}2-\iota\,(3{\cdot}6)$; thus the approximations in (ii.) are not sufficient to calculate either $S_n(\kappa a)$ or $S_n(\kappa_1 a)$ with any accuracy.[‡] In actual fact it proved necessary to use Lord RAYLEIGH'S exact formulæ, equivalent to (4·7) above, and to go as far as $n = 4$ in the series.[§]

§ 6. CASE OF LARGE PERFECTLY CONDUCTING SPHERES.

Before proceeding to the final formulæ, it will be convenient to state certain results given by MACDONALD[||] for the values of the functions $S_n(z)$, $E_n(z)$, when both n and z are large.

---

[*] For the case in which $K-1$ is small this conclusion is given by G. W. WALKER, 'Quarterly Journal of Mathematics,' vol. 30, 1899, p. 217. The formulæ given on that page agree with (5·6), when $K-1$ is small; but the more general formulæ on the preceding page do not agree with (5·6) completely. I have not succeeded in tracing the discrepancy on account of the fact that G. W. WALKER has omitted some of the details of his preliminary calculations.

[†] E. T. PARIS, 'Phil. Mag.,' vol. 30 (Ser. 6), 1915, p. 459.

[‡] To obtain an accuracy of 1 per cent. in $S_1(z)$ by retaining two terms of the series only, it must be supposed that $|z|$ does not exceed 1·3.

[§] E. T. PARIS, *loc. cit.*, p. 472.

[||] 'Phil. Trans.,' vol. 210, A, 1910, p. 134: the formulæ are due to L. LORENZ originally. A very interesting method of deriving the results is given by DEBYE ('Math. Annalen,' Bd. 67, 1909, p. 535).

Provided that $z-(n+\frac{1}{2})$ is of an order higher than $z^{1/3}$, the formulæ are

$$(6\cdot1)\quad \left\{\begin{array}{ll} & E_n(z) = Re^{-\iota\psi}, \qquad S_n(z) = R\sin\psi, \\ \text{where} & \\ & R^2 = 1/\sin\alpha, \qquad \psi = z\sin\alpha+\tfrac{1}{4}\pi-(n+\tfrac{1}{2})\alpha, \\ \text{and} & \\ & \cos\alpha = (n+\tfrac{1}{2})/z. \end{array}\right.$$

We shall need also the corresponding formulæ for $S'_n(z)$ and $E'_n(z)$; it will be seen that

$$\frac{E'_n(z)}{E_n(z)} = \frac{1}{R}\frac{dR}{dz}-\iota\frac{d\psi}{dz} = -\frac{1}{R^2}\left(\frac{1}{2}\frac{\cos\alpha}{\sin^2\alpha}\frac{d\alpha}{dz}+\iota\right),$$

because (3·86) gives

$$R^2\frac{d\psi}{dz} = 1.$$

Hence

$$(6\cdot2)\quad \left\{\begin{array}{ll} & \dfrac{E'_n(z)}{E_n(z)} = -\dfrac{1}{R^2}(\tan\chi+\iota) = -\dfrac{\iota}{R^2}\dfrac{e^{-\iota\chi}}{\cos\chi}, \\ \text{where*} & \\ & \tan\chi = \dfrac{1}{2}\dfrac{\cos\alpha}{\sin^2\alpha}\dfrac{d\alpha}{dz} = \dfrac{1}{2}\dfrac{\cos^2\alpha}{z\sin^3\alpha}. \end{array}\right.$$

Similarly, we find that

$$(6\cdot3)\qquad \frac{S'_n(z)}{S_n(z)} = \frac{1}{R}\frac{dR}{dz}+\cot\psi\frac{d\psi}{dz} = \frac{1}{R^2}(-\tan\chi+\cot\psi) = \frac{1}{R^2}\frac{\cos(\psi+\chi)}{\sin\psi\cos\chi}.$$

The formulæ to be used finally are those for A_n and C_n, given in (4·71), thus we take

$$A_n = -\frac{S'_n(z)}{E'_n(z)}, \qquad C_n = -\frac{S_n(z)}{E_n(z)},$$

where z now denotes κa. It follows from (6·1) above that

$$C_n = -\sin\psi e^{+\iota\psi} = +\tfrac{1}{2}\iota(e^{2\iota\psi}-1),$$

and using (6·2) and (6·3) we see that

$$A_n = -\iota e^{\iota(\psi+\chi)}\cos(\psi+\chi) = -\tfrac{1}{2}\iota\{e^{2\iota(\psi+\chi)}+1\}.$$

It is now an easy matter to write down an approximation to the functions M and N defined in (4·8), provided that θ is not near to 0 or π. Under these conditions we can take the approximate value

$$P_n(\cos\theta) = \sqrt{\left(\frac{2}{n\pi\sin\theta}\right)}\cos\{(n+\tfrac{1}{2})\theta-\tfrac{1}{4}\pi\},$$

---

* Under our conditions α is not near to zero and z is large, so that χ is small.

giving a corresponding approximation

$$\mathrm{P}'_n(\cos\theta) = \frac{2n+1}{\sin\theta}\frac{\sin\{(n+\frac{1}{2})\theta-\frac{1}{4}\pi\}}{\sqrt{(2n\pi\sin\theta)}}.$$

Thus

$$\sin\theta\frac{2n+1}{n(n+1)}\mathrm{A}_n\mathrm{P}'_n(\cos\theta) = -\frac{1}{\sqrt{(2n\pi\sin\theta)}}\{e^{2\iota(\psi+\chi)}+1\}(e^{\iota\omega}-e^{-\iota\omega}),$$

where

$$\omega = (n+\tfrac{1}{2})\theta-\tfrac{1}{4}\pi,$$

and we have replaced $(n+\frac{1}{2})^2/\{n(n+1)\}$ by unity, because n is large.

Thus, on putting $\chi = 0$, we find that (4·8) gives the approximation

$$\text{(6·4)} \qquad \mathrm{M} = -\cos\phi\frac{e^{-\iota\kappa r}}{\kappa r}\Sigma\frac{e^{\iota n\pi}}{\sqrt{(2n\pi\sin\theta)}}(e^{2\iota\psi}+1)(e^{\iota\omega}-e^{-\iota\omega}).$$

Similarly, we get the formula

$$\text{(6·5)} \qquad \mathrm{N} = +\sin\phi\frac{e^{-\iota\kappa r}}{\kappa r}\Sigma\frac{e^{\iota n\pi}}{\sqrt{(2n\pi\sin\theta)}}(e^{2\iota\psi}-1)(e^{\iota\omega}-e^{-\iota\omega}).$$

With series of this type, the leading part is found by making the index of the exponential stationary (regarded as a function of n). Now in both M and N there is one index only, $\sigma = 2\psi+n\pi-\omega$, which can be stationary: and the condition is

$$\frac{d\sigma}{dn} = 2\frac{d\psi}{dn}+\pi-\theta = 0.$$

Now, from (6·1)

$$\frac{d\psi}{dn} = \{z\cos\alpha-(n+\tfrac{1}{2})\}\frac{d\alpha}{dn}-\alpha = -\alpha,$$

and so the leading terms in (6·4) and (6·5) arise from taking

$$2\alpha = \pi-\theta, \quad \text{or} \quad n+\tfrac{1}{2} = z\sin\tfrac{1}{2}\theta.$$

The corresponding value of the index σ is then

$$\begin{aligned}\sigma_0 = 2\psi+n\pi-\omega &= 2z\sin\alpha+\tfrac{1}{2}\pi-(2n+1)\alpha+n\pi-(n+\tfrac{1}{2})\theta+\tfrac{1}{4}\pi,\\ &= 2z\cos\tfrac{1}{2}\theta+\tfrac{1}{4}\pi.\end{aligned}$$

To determine the form of the index near to this special value of n we take

$$\frac{d^2\sigma}{dn^2} = -2\frac{d\alpha}{dn} = \frac{2}{z\sin\alpha} = \frac{2}{z\cos\frac{1}{2}\theta},$$

and then we find the approximate formulæ

$$\text{(6·51)} \quad \left\{\begin{array}{ll} & \sigma = \sigma_0+\dfrac{(n-n_0)^2}{z\cos\frac{1}{2}\theta}, \\ \text{where} & \\ & n_0+\tfrac{1}{2} = z\sin\tfrac{1}{2}\theta. \end{array}\right.$$

The leading parts of M, N are accordingly given by the approximations

$$(6{\cdot}6) \qquad \mathrm{M} = +\cos\phi\frac{e^{-\iota\kappa r}}{\kappa r}\mathrm{Q}, \qquad \mathrm{N} = -\sin\phi\frac{e^{-\iota\kappa r}}{\kappa r}\mathrm{Q},$$

where

$$\mathrm{Q} = \Sigma\frac{e^{\iota\sigma}}{\sqrt{(2n\pi\sin\theta)}},$$

and σ has the value given in (6·51).

The value of Q is approximately equal to the integral

$$(6{\cdot}61) \qquad \int_{-\infty}^{\infty}\frac{e^{\iota\sigma}\,dx}{\sqrt{(2n_0\pi\sin\theta)}},$$

where

$$n = n_0+x, \qquad \sigma = \sigma_0 + x^2/(z\cos\tfrac{1}{2}\theta).$$

Thus, approximately

$$\mathrm{Q} = \frac{e^{\iota\sigma_0}}{\sqrt{(2n_0\pi\sin\theta)}}\sqrt{\left(\frac{\pi z\cos\frac{1}{2}\theta}{-\iota}\right)}$$

$$= \frac{e^{\iota(\sigma_0+\frac{1}{4}\pi)}}{2\sin\frac{1}{2}\theta} = \frac{\iota e^{\iota(2z\cos\frac{1}{2}\theta)}}{2\sin\frac{1}{2}\theta}.$$

Accordingly, to the same degree of approximation, we can take

$$(6{\cdot}62) \qquad \frac{d\mathrm{Q}}{d\theta} = \tfrac{1}{2}ze^{2\iota z\cos\frac{1}{2}\theta}.$$

Then the components of force are given, as in (4·9), by

$$(6{\cdot}63) \qquad \left\{\begin{aligned} \mathrm{Y} &= +c\gamma = \frac{\partial\mathrm{M}}{\partial\theta} - \frac{1}{\sin\theta}\frac{\partial\mathrm{N}}{\partial\phi} = \cos\phi\frac{e^{-\iota\kappa r}}{\kappa r}\frac{d\mathrm{Q}}{d\theta}, \\ \mathrm{Z} &= -c\beta = \frac{1}{\sin\theta}\frac{\partial\mathrm{M}}{\partial\phi} + \frac{\partial\mathrm{N}}{\partial\theta} = -\sin\phi\frac{e^{-\iota\kappa r}}{\kappa r}\frac{d\mathrm{Q}}{d\theta}, \end{aligned}\right.$$

where differential coefficients with respect to ϕ are small compared with those with respect to θ, and so have been rejected. Thus, using (6·62), we have the approximations to the forces in the scattered waves

$$(6{\cdot}7) \qquad \left\{\begin{aligned} \mathrm{Y} &= +c\gamma = \cos\phi\frac{a}{2r}e^{2\iota\kappa a\cos\frac{1}{2}\theta-\iota\kappa r}, \\ \mathrm{Z} &= -c\beta = -\sin\phi\frac{a}{2r}e^{2\iota\kappa a\cos\frac{1}{2}\theta-\iota\kappa r}, \end{aligned}\right.$$

assuming that θ is neither near to 0 nor to π.

When θ is small, the approximation to $\mathrm{P}_n(\cos\theta)$ must be taken as

$$\mathrm{P}_n(\cos\theta) = \mathrm{J}_0\{(2n+1)\sin\tfrac{1}{2}\theta\},$$

and so

$$\sin\theta\mathrm{P}'_n(\cos\theta) = (n+\tfrac{1}{2})\cos\tfrac{1}{2}\theta\,\mathrm{J}_1\{(2n+1)\sin\tfrac{1}{2}\theta\}.$$

Thus we now write

$$\sin\theta\,\frac{2n+1}{n(n+1)}\,A_n P'_n(\cos\theta) = -\iota\,(e^{2\iota\psi}+1)\cos\tfrac{1}{2}\theta\,J_1\,\{(2n+1)\sin\tfrac{1}{2}\theta\},$$

and, proceeding as before, we are led to the conclusion that α is near to $\frac{1}{2}\pi$; thus the value of n/z is small, in the parts of the series which contribute the principal part of the sum. Then we can replace (6·1) by the approximate formulæ

$$(6{\cdot}8)\quad \left\{\begin{aligned} &\alpha = \tfrac{1}{2}\pi-(n+\tfrac{1}{2})/z \\ \text{and}\quad &\psi = z+\tfrac{1}{4}\pi-(n+\tfrac{1}{2})(\tfrac{1}{2}\pi)+\tfrac{1}{2}(n+\tfrac{1}{2})^2/z. \end{aligned}\right.$$

Thus

$$2\psi+n\pi = 2z+(n+\tfrac{1}{2})^2/z.$$

Hence the approximation corresponding to (6·4) is now

$$(6{\cdot}81)\qquad M = -\iota\cos\tfrac{1}{2}\theta\cos\phi\,\frac{e^{-\iota\kappa r}}{\kappa r}\,\Sigma\, e^{2\iota z+\iota(n+\frac{1}{2})^2/z}J_1\,\{2n+1)\sin\tfrac{1}{2}\theta\}.$$

In like manner the value of N is found to differ from (6·81) only in having $+\sin\phi$ as a factor instead of $-\cos\phi$.

In the series (6·81) the value of n may be supposed to vary from 0 to ∞; and so we obtain the principal part of the sum by using the integral

$$(6{\cdot}82)\qquad e^{2\iota z}\int_0^\infty e^{\iota\zeta^2/z}J_1\,(2\zeta\sin\tfrac{1}{2}\theta)\,d\zeta;$$

and when $\sin\frac{1}{2}\theta$ is very small the value of (6·82) is approximately equal to

$$e^{2\iota z}\,(\tfrac{1}{2}\iota z\sin\tfrac{1}{2}\theta).$$

Thus

$$(6{\cdot}83)\quad \left\{\begin{aligned} &M = +\tfrac{1}{4}z\sin\theta\cos\phi\, e^{2\iota z}\,\frac{e^{-\iota\kappa r}}{\kappa r}, \\ \text{and}\quad &N = -\tfrac{1}{4}z\sin\theta\sin\phi\, e^{2\iota z}\,\frac{e^{-\iota\kappa r}}{\kappa r}. \end{aligned}\right.$$

Accordingly the components of force are now found to be

$$(6{\cdot}9)\quad \left\{\begin{aligned} Y &= +c\gamma = \frac{\partial M}{\partial\theta}-\frac{1}{\sin\theta}\frac{\partial N}{\partial\phi} = \tfrac{1}{2}z\cos\phi\,\frac{e^{2\iota z-\iota\kappa r}}{\kappa r}, \\ Z &= -c\beta = \frac{1}{\sin\theta}\frac{\partial M}{\partial\phi}+\frac{\partial N}{\partial\theta} = -\tfrac{1}{2}z\sin\phi\,\frac{e^{2\iota z-\iota\kappa r}}{\kappa r}. \end{aligned}\right.$$

These results in (6·9) are precisely the same as would be found by writing $\theta = 0$ in the approximations (6·7), and *accordingly the formulæ* (6·7) *do remain valid right up to the axis* $\theta = 0$.

When θ is nearly equal to π, the calculation on the present lines becomes more difficult,* and we shall accordingly obtain the corresponding approximation by a different process in the next section (§ 7).

It appears that for the special value $\kappa a = 10$, the formulæ (6·7) do give the forces with a fair degree of accuracy up to an angle $\theta = \frac{2}{3}\pi$; the approximation in fact appears to be better than might have been expected. [See p. 176 above.]

§ 7. ALTERNATIVE METHOD, APPLICABLE TO ANY CONDUCTOR WHOSE DIMENSIONS ARE LARGE COMPARED WITH THE WAVE-LENGTH.

It follows at once from GREEN'S theorem that if u, v are solutions of the equations

$$\Delta^2 u + \kappa^2 u = 0, \qquad \Delta^2 v + \kappa^2 v = 0,$$

at points within a closed simple surface S, then

$$\int \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) d\mathrm{S} = 0, \tag{7·1}$$

where the integral is taken over the surface S ($d\nu$ being the element of outward normal), and it is supposed that u, v are both free from singularities in the interior of S.

Similarly if u has no singularities and v is a solution which behaves like $e^{-\iota\kappa \mathrm{R}}/\mathrm{R}$ near a particular point P (R denoting the distance measured from P), we see that

$$\int \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) d\mathrm{S} = -4\pi u_{\mathrm{P}}, \tag{7·11}$$

provided that P is inside the surface S.

Equations of similar forms apply when the space considered is *outside* the surface S; but then the sign of the last equation (7·11) is reversed, giving

$$\int \left(u \frac{\partial v}{\partial \nu} - v \frac{\partial u}{\partial \nu} \right) d\mathrm{S} = +4\pi u_{\mathrm{P}}; \tag{7·12}$$

it is then necessary to assume also that at infinity, u, v both correspond to divergent waves (unless it is known that u tends to zero more rapidly than $1/r$).

---

* Compare MACDONALD, *loc. cit.*, pp. 120–122. The cause of the difficulty is to be found in the fact that now the stationary value may be expected to arise from values of n for which α is small. Then n is nearly equal to z, and in all such cases more complicated analysis is inevitable. In fact the approximations to $\mathrm{S}_n(z)$ and $\mathrm{E}_n(z)$ require to be modified by different formulæ corresponding to the cases $n > z$, $n < z$ and to the cases in which $|n - z|$ is of order $z^{1/3}$ or of lower order.

Let us now consider the problem of waves incident from some source (or sourc and reflected from the surface S. Let u denote any Cartesian component of force the incident wave, and let the point P be outside S. Then (7·1) applies, and so

$$\int\left(u\frac{\partial v}{\partial\nu}-v\frac{\partial u}{\partial\nu}\right)dS = 0,$$

because u has no singularity inside S.

If u' denotes the corresponding component of force in the reflected wave, we ha from (7·12)

$$\int\left(u'\frac{\partial v}{\partial\nu}-v\frac{\partial u'}{\partial\nu}\right)dS = 4\pi u'_{\mathrm{P}};$$

because u' has no singularity outside S (and u' will correspond to a divergent w at infinity).

By addition we have the result

$$(7{\cdot}13)\qquad 4\pi u'_{\mathrm{P}} = \int\left\{(u+u')\frac{\partial v}{\partial\nu}-v\frac{\partial}{\partial\nu}(u+u')\right\}dS.$$

where v is taken to be $e^{-\iota\kappa\mathrm{R}}/\mathrm{R}$.

Now $u+u' = w$ gives the corresponding component of force in the complete wa and this accordingly satisfies certain known relations at the surface S (the ex form depending on the physical properties of S). It must, however, be clea understood that we cannot usually obtain both w and $\frac{\partial w}{\partial\nu}$ by any simple methods, a more than the analogous problems of electrostatics can be solved by a mere app to GREEN'S Theorem.

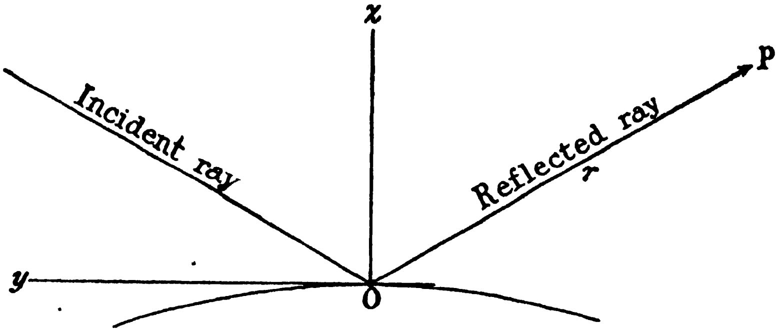
However, we can obtain *an approximate solution, suitable to the problem of sh wave-lengths*, by assuming that near the reflecting surface, the character of u' can determined from that of u by the rules of elementary geometrical optics. Thus treat the reflected wave as derived from the incident by simple reflexion *in tangent-plane* at the point of incidence. Making this hypothesis it is an easy mat to construct both w and $\frac{\partial w}{\partial\nu}$, when the form of u is given.

It will be noticed that we shall have $w = 0$, $\frac{\partial w}{\partial\nu} = 0$ at all points within t geometrical shadow; and so the final integral (7·13) extends only over the illuminat side of the surface S.

Suppose now that we consider electric waves incident on a simple con conducting surface; and take an origin O on the surface such that OP is t reflected ray (in the sense of geometrical optics). Take the plane of incidence as t plane of yz, and the normal at O as the axis of z.

Then in the immediate neighbourhood of O we can represent the incident wave by the components of electric force

$$(\mathrm{A},\ \mathrm{B},\ \mathrm{C})\,e^{\iota\kappa(my+nz)}$$



where $m^2+n^2 = 1$, so that m is the cosine and n the sine of the angle of incidence. We have also the relation

$$\mathrm{B}m+\mathrm{C}n = 0$$

because the electric force is perpendicular to the incident ray.

The corresponding reflected wave has the components

$$(\mathrm{A}',\ \mathrm{B}',\ \mathrm{C}')\,e^{\iota\kappa(my-nz)}$$

where

$$\mathrm{A}+\mathrm{A}' = 0, \qquad \mathrm{B}+\mathrm{B}' = 0,$$

and

$$\mathrm{B}'m-\mathrm{C}'n = 0,$$

so that

$$\mathrm{C}'-\mathrm{C} = 0\,;$$

these results follow by making the tangential components of force zero on the tangent-plane $z = 0$ (instead of at the surface).

It is now clear that, at the point O, the components of the total force will be

$$(7{\cdot}2) \qquad (0,\ 0,\ 2\mathrm{C}),$$

and the normal differential coefficients of the total force will be

$$(7{\cdot}21) \qquad \iota\kappa n\,(\mathrm{A}-\mathrm{A}',\ \mathrm{B}-\mathrm{B}',\ \mathrm{C}-\mathrm{C}') = 2\iota\kappa n\,(\mathrm{A},\ \mathrm{B},\ 0).$$

We shall now insert these values for w and $\frac{\partial w}{\partial \nu}$ in the general formula (7·13), including also the factor $e^{\iota\kappa(my+nz)}$, on account of phase-differences at points near to O. Since reflexion takes place only from the immediate neighbourhood of O, the error introduced by this simplification will be small.

Since the co-ordinates of P are $(0, -mr, nr)$, the value of R is given by

$$\begin{aligned} R^2 &= x^2+(y+mr)^2+(z-nr)^2, \\ &= r^2+2r(my-nz)+x^2+y^2+z^2, \end{aligned}$$

where (x, y, z) is a point on the surface near to O. Thus, when r is very large in comparison with the dimensions of the surface (as we assumed in the previous investigations, §§ 4–6), we can use the approximate formula

$$R = r+my-nz. \tag{7·3}$$

Thus we can write in (7·13)

$$v = \frac{e^{-\iota\kappa(r+my-nz)}}{r} = v_0 e^{-\iota\kappa(my-nz)},$$

if v_0 is the value of v at O. Then the most important term in $\frac{\partial v}{\partial \nu}$ is seen to be

$$\frac{\partial v}{\partial z} = \iota\kappa n v_0 e^{-\iota\kappa(my-nz)}.$$

Accordingly the components of electric force in the reflected wave will be given by the approximation

$$(7·4) \quad \left\{ \begin{aligned} X &= -\frac{v_0}{4\pi}\int(2\iota\kappa n A)\, e^{2\iota\kappa nz}\, dS, \\ Y &= -\frac{v_0}{4\pi}\int(2\iota\kappa n B)\, e^{2\iota\kappa nz}\, dS, \\ Z &= +\frac{v_0}{4\pi}\int(2\iota\kappa n C)\, e^{2\iota\kappa nz}\, dS. \end{aligned} \right.$$

To evaluate the integrals in (7·4) we must write out the equation to the surface in the approximate form

$$2z = -(\alpha x^2+2\beta xy+\gamma y^2).$$

Then*

$$\int e^{-\iota\kappa n(\alpha x^2+2\beta xy+\gamma y^2)}\, dS = \frac{\pi}{\iota\kappa n}\frac{1}{\sqrt{(\alpha\gamma-\beta^2)}}.$$

Now $\alpha\gamma-\beta^2$ is the absolute (or Gaussian) curvature of the surface at the point O; and since the surface is supposed convex, we represent this curvature by $1/\rho^2$.

---

* This is most easily found by taking the integral as $\lim\limits_{\theta\to 0}\int e^{-(\theta+\iota\kappa)n(\alpha x^2+2\beta xy+\gamma y^2)}\, dS$.

Thus we get

$$(7{\cdot}41)\qquad \iota\kappa n\int e^{2\iota\kappa n z}\,d\mathrm{S} = \pi\rho.$$

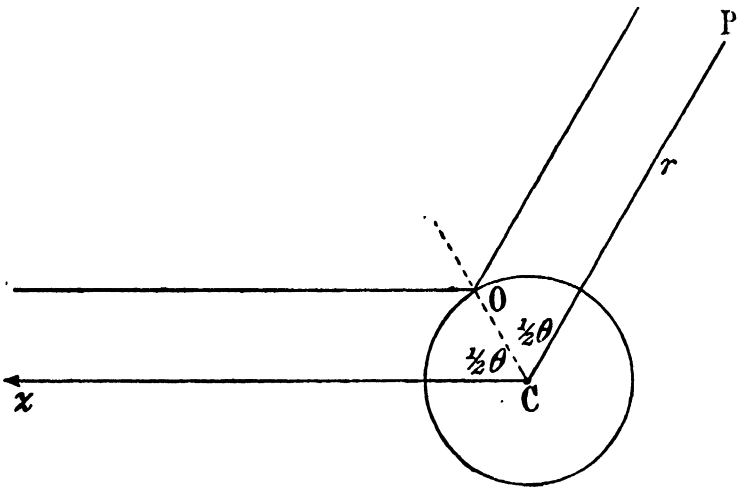
Hence, using (7·4), we see that the principal part of the reflected wave is given by

$$(7{\cdot}5)\qquad (\mathrm{X},\ \mathrm{Y},\ \mathrm{Z}) = (-\mathrm{A},\ -\mathrm{B},\ +\mathrm{C})\frac{\rho}{2r}e^{-\iota\kappa r}.$$

In order to interpret (7·5) for any axes of co-ordinates, we need only notice that (−A, −B, +C) represents a force numerically equal to the force in the incident wave; and that the new force is perpendicular to the reflected ray, arranged in such a way that the tangential components are opposite to those in the incident wave.

We can apply the formula (7·5) to the problem of § 6 at once; clearly $\rho = a$.

The point of incidence corresponding to the scattered wave (θ, ϕ) is given by $(\frac{1}{2}\theta, \phi)$.



Then the incident wave at O has the components of electric force

$$-\cos\phi e^{\iota\kappa a\cos\frac{1}{2}\theta} \text{ in the plane of incidence,}$$

and

$$+\sin\phi e^{\iota\kappa a\cos\frac{1}{2}\theta} \text{ perpendicular to the plane of incidence.}$$

Further, the r of formula (7·5) is measured from O; to compare with § 6, we take r to be the distance CP, measured from the centre of the sphere. Thus we are to replace r in (7·5) by $r - a\cos\frac{1}{2}\theta$. Accordingly the components of force at P, in the reflected wave, are

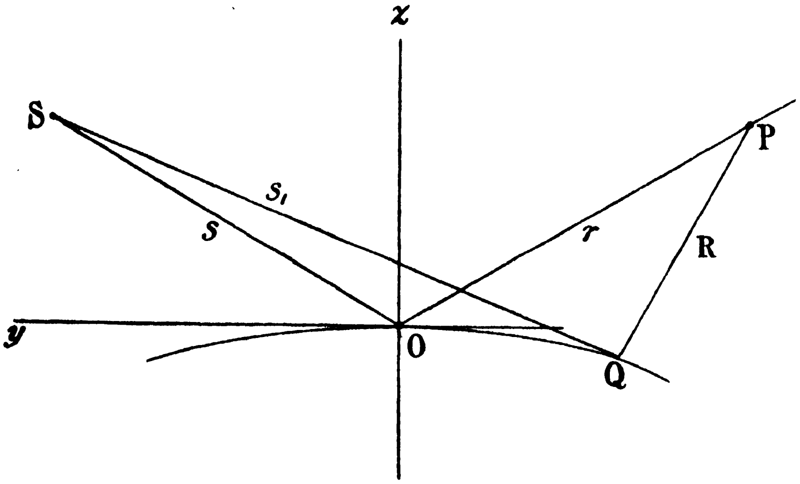
$$+\frac{a}{2r}\cos\phi e^{2\iota\kappa a\cos\frac{1}{2}\theta-\iota\kappa r} \text{ perpendicular to } r \text{ in the plane ZCP,}$$

and

$$-\frac{a}{2r}\sin\phi e^{2\iota\kappa a\cos\frac{1}{2}\theta-\iota\kappa r} \text{ perpendicular to the plane ZCP.}$$

These results agree with (6·7) and (6·9) of § 6 above.

It is easy to modify the general formulæ (7·5) so as to cover the case of waves incident from a point source (say at distance s from the point of incidence).



Then

$$s_1^2 = x^2 + (ms - y)^2 + (ns - z)^2$$
$$= s^2 - 2s(my + nz) + x^2 + y^2 + z^2.$$

Thus with the usual approximation of geometrical optics

$$s_1 = s - (my + nz) + \frac{1}{2s}\{x^2 + (ny - mz)^2\}.$$

Similarly

$$\mathrm{R} = r + my - nz + \frac{1}{2r}\{x^2 + (ny + mz)^2\}.$$

Now z is of the second order in comparison with x, y; and so we can write

$$s_1 + \mathrm{R} = s + r - n(\alpha x^2 + 2\beta xy + \gamma y^2) + \left(\frac{1}{2r} + \frac{1}{2s}\right)(x^2 + n^2 y^2).$$

Thus here we find

$$\int e^{-\iota\kappa(s_1+\mathrm{R})}\, d\mathrm{S} = e^{-\iota\kappa(s+r)}\frac{\pi}{\iota\kappa n\sigma}$$

where

$$\sigma^2 = \left\{\tfrac{1}{2}\left(\frac{1}{r} + \frac{1}{s}\right) - n\alpha\right\}\left\{\tfrac{1}{2}\left(\frac{1}{r} + \frac{1}{s}\right) - \frac{\gamma}{n}\right\} - \beta^2$$
$$= \tfrac{1}{4}\left(\frac{1}{r} - \frac{1}{r_1}\right)\left(\frac{1}{r} - \frac{1}{r_2}\right),$$

r_1, r_2 being the distances of the focal lines (of geometrical optics) from the point of incidence.

Thus now the principal parts of the reflected wave are given by

$$(-\mathrm{A},\ -\mathrm{B},\ +\mathrm{C})\,e^{-\iota\kappa r}\Big/\left\{\left(1 - \frac{r}{r_1}\right)\left(1 - \frac{r}{r_2}\right)\right\}^{1/2};$$

assuming that r is not close either to r_1, or to r_2.

The results of the foregoing analysis depend on the tacit assumption that n is not zero; and as a consequence the character of the approximations will change when n

is small. That is, near the edge of the shadow, in the ordinary phrase of geometrical optics.

In the application to the sphere, the region excluded by this condition corresponds to values of θ nearly equal to π; and in this region the specification of the scattered wave by means of Cartesian co-ordinates seems simplest.

We consider then the incident wave as specified in Cartesian form by

$$(7\cdot 6) \qquad X = -e^{\iota\kappa z}, \qquad Y = 0, \qquad Z = 0,$$

and consider the approximation to the reflected wave incident at the point (al, am, an) on the sphere for which the direction-cosines of the normal are l, m, n. The expressions will be of the form

$$(7\cdot 61) \qquad (X', Y', Z') = (A', B', C')e^{\iota\kappa\zeta}$$

where (treating the tangent-plane as the reflecting surface)

$$\zeta = z+p\,(a-lx-my-nz)$$

and p is determined by

$$(-lp)^2+(-mp)^2+(1-np)^2 = 1.$$

Thus

$$p = 2n.$$

Further, the resultant of (X, Y, Z) and (X′, Y′, Z′) at the point of incidence must be along the normal; and so

$$\frac{A'-1}{l} = \frac{B'}{m} = \frac{C'}{n}.$$

Also (A′, B′, C′) is perpendicular to the reflected ray; and so

$$A'(-2nl)+B'(-2nm)+C'(1-2n^2) = 0.$$

Hence

$$(7\cdot 62) \quad \left\{ \begin{array}{l} \dfrac{A'-1}{l} = \dfrac{B'}{m} = \dfrac{C'}{n} = \dfrac{2nl}{n-2n} = -2l, \\ \text{or} \\ A' = 1-2l^2, \quad B' = -2lm, \quad C' = -2ln. \end{array} \right.$$

The components X+X′, Y+Y′, Z+Z′ at the point of incidence are accordingly equal to

$$(7\cdot 63) \qquad -2\,(l^2, lm, ln)\,e^{\iota\kappa an}.$$

We have still to evaluate the normal differential coefficients, which are found to be

$$(7\cdot 64) \qquad \iota\kappa n\,(-1-A', -B', -C')\,e^{\iota\kappa an} = 2\iota\kappa n\,(l^2-1, lm, ln)\,e^{\iota\kappa an}.$$

The value of R is now seen to be given by

$$R^2 = (x-al)^2+(y-am)^2+(z-an)^2 = r^2-2a\,(lx+my+nz)+a^2.$$

Thus, when we regard a/r as small, we may take

$$(7{\cdot}65)\quad \left\{\begin{array}{ll} & \mathrm{R} = r-a\,(l\sin\theta\cos\phi+m\sin\theta\sin\phi+n\cos\theta), \\ \text{and} & \\ & \dfrac{\partial \mathrm{R}}{\partial \nu} = -\,(l\sin\theta\cos\phi+m\sin\theta\sin\phi+n\cos\theta). \end{array}\right.$$

Now, in applying (7·13), $\frac{\partial \mathrm{R}}{\partial \nu}$ occurs only in the coefficients and not in the exponential index: thus we can get the first approximation by putting $\theta = \pi$ in $\frac{\partial \mathrm{R}}{\partial \nu}$; this gives the value

$$(7{\cdot}66)\qquad \frac{\partial \mathrm{R}}{\partial \nu} = n.$$

Thus, to our degree of accuracy

$$\begin{aligned} \frac{1}{4\pi}\left(w\frac{\partial v}{\partial \nu} - v\frac{\partial w}{\partial \nu}\right) &= -\frac{v}{4\pi}\left(\iota\kappa w\frac{\partial \mathrm{R}}{\partial \nu} + \frac{\partial w}{\partial \nu}\right), \\ &= -\frac{v}{4\pi}\left(\iota\kappa n w + \frac{\partial w}{\partial \nu}\right). \end{aligned}$$

We can now substitute for w and $\frac{\partial w}{\partial \nu}$ the values given by (7·63) and (7·64): it will be seen that the components parallel to y, z give zero (to this order), and that the component parallel to x gives

$$\frac{\iota\kappa n}{2\pi}\,\frac{e^{\iota\kappa(an-\mathrm{R})}}{r}.$$

Accordingly the reflected wave is given by

$$(7{\cdot}67)\qquad \mathrm{X} = \frac{\iota\kappa}{2\pi r}\int n\,d\mathrm{S}\,e^{\iota\kappa(an-\mathrm{R})},$$

where R is found from (7·65) and the integral extends over the positive hemisphere. For the purpose of integration we write

$$l = \sin\theta'\cos\phi', \qquad m = \sin\theta'\sin\phi', \qquad n = \cos\theta'.$$

Then in (7·67) we have

$$an-\mathrm{R} = -r+a\,\{(1+\cos\theta)\cos\theta'+\sin\theta\sin\theta'\cos(\phi-\phi')\}.$$

The integration of (7·67) with respect to ϕ' can be carried out at once, because

$$(7{\cdot}68)\qquad \int_0^{2\pi} e^{\iota\kappa a\sin\theta\sin\theta'\cos(\phi-\phi')}\,d\phi' = 2\pi \mathrm{J}_0(\kappa a\sin\theta\sin\theta').$$

Our result accordingly becomes

$$(7{\cdot}69) \qquad \mathrm{X} = \frac{\iota\kappa a^2}{r} e^{-\iota\kappa r} \int_0^{\frac{1}{2}\pi} e^{\iota\kappa a(1+\cos\theta)\cos\theta'} \mathrm{J}_0(\kappa a \sin\theta \sin\theta') \sin\theta' \cos\theta' \, d\theta'.$$

An exact evaluation of the integral would be troublesome; but it can be transformed by integration by parts. This process leads to a series of which the first term is

$$(7{\cdot}7) \qquad \mathrm{X} = \frac{\iota\kappa a^2}{r} e^{-\iota\kappa r} \frac{\mathrm{J}_1(\kappa a \sin\theta)}{\kappa a \sin\theta}.$$

The formula (7·7) will represent (7·69) sufficiently accurately if $(1+\cos\theta)/\sin\theta = \cot\frac{1}{2}\theta$ is regarded as small; and this is the correct assumption here, since θ is supposed to be nearly equal to π.

In the special cases $\kappa a = 9, 10$ it appears that the approximation (7·7) represents the scattered wave sufficiently well within a cone extending to about 10° from the axis.

VII. *On the Fundamental Formulations of Electrodynamics.*

By G. H. LIVENS, *University of Manchester.*

Communicated by W. M. HICKS, *F.R.S.*

Received March 14,—Read June 26, 1919.

1. MODERN electrical theory based on MAXWELL'S concept of an æthereal displacement current, is generally regarded as being sufficiently complete in itself to cover all actions so far revealed to us, if we exclude those intra-atomic phenomena which probably involve some additional but not necessarily inconsistent action in their working. There, still, however exists a good deal of uncertainty as to the actual results of the development of this theory in certain directions, and no account has yet been taken of the great degree of latitude allowed by it in its simplest and most general form. For example, in most presentations of the theory of energy streaming in the electromagnetic field the discussion is given in a way which might lead one to believe that POYNTING'S form* of the theory is the only one conceivable. A single alternative has on one occasion† been suggested, but rather as an improvement on POYNTING'S form than as an indication of its uncertainty. Whilst it cannot be denied that POYNTING'S theory is probably the most appropriate one yet formulated, yet it must be recognised that there are an infinite number of fundamentally different forms each of which is itself perfectly consistent with MAXWELL'S theory as expressed in his differential equations of electromagnetic interaction.

Again, but now we are on a different plane, it has usually been stated that MAXWELL'S theory is not of sufficient generality to cover the cases where there exists the complication of non-linear induction in ferromagnetic media.‡ This view appears to have originated with the idea that the *magnetic force* is the fundamental æthereal vector of the magnetic field, whereas, as a matter of fact, the only consistent view § of the energy relations of such a field leads to the conclusion that the *magnetic induction* is the true æthereal vector, the magnetic force being an auxiliary vector

---

* 'Phil. Trans.,' A, vol. 175 (1884).

† MACDONALD, 'Electric Waves,' Chs. IV., V., VIII.

‡ This is the view of practically all Continental writers on this subject.

§ *Cf.* 'Roy. Soc. Proc.,' A, vol. 93, p. 20 (1916).

[Published March 17, 1920.

derived in the process of averaging the minute current whirls into their effective representation as a distribution of magnetic polarity.*

Further, the expression for the forcive of electromagnetic origin acting on the elements of a polarised medium still seems to be the subject of some doubt. MAXWELL† derives an expression in the magnetic case by statical considerations based on the method of energy, and then seems to regard it as generally valid under all circumstances. Objection has however been taken to MAXWELL'S expression by certain writers who, basing themselves on the presumed analogy between the dielectric and magnetic cases, prefer a form of expression differing from MAXWELL'S by a quantity which vanishes in the statical case considered but which is of fundamental importance in the derived problem of reducing the general forcive of electrodynamic origin to a representation by means of an imposed stress system. It appears in fact that the presence of this extra part in MAXWELL'S expression is effective in securing the ordinary expression for the subsidiary term arising in the induction, which has given rise to the conception of electromagnetic momentum, on account of its being a perfect time differential. In the alternative form of the theory the perfect time differential is not secured so that the idea of electromagnetic momentum is lost.‡ In his edition of MAXWELL'S treatise, J. J. THOMSON adds a note attempting to justify MAXWELL'S form of the expression, but his discussion can easily be shown to be erroneous, for he fails to distinguish between the true and complete currents of the theory, the latter containing a constituent, viz., the rate of change of æthereal displacement, which is not affected by the magnetic part of the complete electromagnetic forcive; nevertheless, the later discussions of the question from the point of view of the theory of electrons have confirmed MAXWELL'S original expression for the magnetic forcive, but they apparently still give the alternative expression for the dielectric case.

It was with the view to clearing up these and certain other difficulties that the present discussion was undertaken, the object aimed at being the formulation of a complete and precise statement of the theory in the only form in which it is logically consistent, then to compare this form with current statements of the theory,§ and finally to exhibit in their true aspects the various derived theories which are included in the general scheme. The original differential theory will be linked up with the subsequent dynamical theories by a discussion in its most general form of the derivation of the fundamental equations based on the principle of Least Action, in the course of which certain inconclusive aspects of this derivation present themselves for consideration.

---

* *Cf.* LARMOR, 'Roy. Soc. Proc.,' vol. 71 (1903), "On the Mechanical and Thermal Relations of the Energy of Magnetisation."

† 'Treatise,' II., Ch. II.

‡ *Cf.* LEATHEM, 'Roy. Soc. Proc.,' A, vol. 89 (1913), p. 34. In this note Mr. LEATHEM attempts to avoid the discrepancy by adding a new term to the force in the elementary polar theory. His only argument in favour of this force is that it overcomes the difficulty, so that it is not very convincing.

§ A complete statement of the fundamental results of the theory so far as they existed up till 1916 is given in my treatise, 'The Theory of Electricity' (Cambridge, 1918). The present paper may be regarded in some measure as a correction and generalisation of the statement there given.

2. A complete statement of MAXWELL'S theory as originally given and in the form which will include most of the recent extensions depends on certain field vectors which first require consistent and independent definition. These vectors:—

(*a*) E, the *electric force*, defined at any point of the field as the vectorial ratio to a small electric charge of the force acting on it when placed at rest in that position. When the point under consideration is inside the matter in the field there exists the possibility of an additional contribution to this force due to local conditions of the matter in the neighbourhood of the point, but for the present we shall disregard any complication of this kind.

(*b*) C, the complete electric current; in the most general case this consists of several distinct parts. Firstly, there is the differential drift of free ions constituting the true conduction current and the material dielectric displacement current; then there is a part due to the convection of charged and electrically polarised media, and finally the æthereal constituent essential and peculiar to MAXWELL'S theory.

It has been definitely established that all but the last constituent of the current are in themselves true movements of electricity, or at least effectively equivalent to such, so that so long as we retain the definite concept of an electrical entity the origin of these different constituents merely remains a matter of kinematics, and they have, in fact, been fully dealt with on this basis.*

(*c*) H, the magnetic force, defined in a similar manner to the electric force, with the aid of the concept of a magnetic pole, but now *without* the possibility of a local contribution due to surrounding magnetic polarity if the point is inside the matter.

It is perhaps as well to emphasise the fact that both the electric and magnetic forces are defined in a theoretical manner which almost excludes the possibility of direct experimental verification. Electromagnetic measurements, particularly as regards the fields inside the matter, are concerned almost entirely with matter in bulk, and it is then only the mechanical or molar parts of these forces that are then under observation, the local parts, if they exist at all, being balanced on the spot by other forces of an origin not at present under review. We have however evidence that these local parts of the forces do exist.

(*d*) B, the magnetic induction, which is defined in the elementary theory in terms of the magnetic force H and the magnetic polarisation intensity I by the relation

$$B = H + 4\pi I,$$

and which is always assumed to be subject to the relation

$$\operatorname{div} B = 0.$$

---

* *Cf.* my 'Theory of Electricity,' p. 363.

In the dynamical theory, however, this vector B appears as the complete æthereal magnetic force,* the part H being merely the mechanically effective part of this force in the material media. This new conception of the magnetic force, which is supported by such phenomena as the Hall effect, &c., where the deviations in ferromagnetic media are proportional to the induction,† is not really inconsistent with the previous definition given under (*c*) above. As there explained, although the vector H is defined as the force on unit magnetic pole, there is still the possibility that there may exist in addition a force ($4\pi I$) not directly detectable in a mechanical experiment.

(*e*) In addition to these fundamental field vectors there are certain others of a subsidiary nature determining the electric and magnetic conditions in the media. One of these, the magnetic polarisation intensity I, has already been introduced, and there is an electric analogue in P, the dielectric polarisation intensity; these two vectors define respectively the effective resultant magnetic and electric bi-polar moments per unit volume of the media.

3. MAXWELL's theory may now be expressed in the statement that in the most general case of interaction between the magnetic and electric fields the four fundamental vectors defining these fields are subject to the two differential vector equations

$$\text{Curl } H = \frac{4\pi C}{c},$$

$$\text{Curl } E = -\frac{1}{c}\frac{dB}{dt},$$

together with the scalar relation defining the electric charge density ρ

$$\text{div } E = 4\pi\rho.$$

The second vectorial equation practically implies that

$$\text{div } B = 0,$$

so that it is hardly necessary to add this as an additional equation.

For the further development of the theory all that is now necessary‡ is the kinematical specification of the current C in terms of the material vectors and an expression for the rate of change of the magnetic force vector at any point of the field. The first of these has the usual form

---

* This was first clearly recognised by LARMOR. *Cf.* his papers "On a Dynamical Theory of the Electric and Luminiferous Medium," 'Phil. Trans.,' 1894–1897, or 'Æther and Matter' (Cambridge, 1900).

† The results are interpreted usually as implying that the deviations are proportional to the polarisation intensity in such cases, but this is equivalent to the statement given.

‡ We are not here concerned with the constitutional equations giving the laws of induction and conduction. These form a separate branch of the subject, whose results are irrelevant to the present discussion.

$$C = \frac{dD}{dt} + \rho\nu_m + C_1,$$

wherein C_1 is the density of the true conduction current, ν_m is the velocity of the material medium at the typical field point, and

$$D = \frac{1}{4\pi}E + P$$

is the total dielectric displacement of MAXWELL'S theory which consists in part of the dielectric polarisation P and in part of an æthereal constituent proportional to the electric force.

The time rate of change of the composite vector D requires careful specification; it consists in the main of the terms

$$\frac{1}{4\pi}\frac{dE}{dt} + \frac{dP}{dt},$$

but when the dielectric media are in motion there is in addition a term arising on account of the convection of the polarisation. This term has been shown* to be equal to

$$\text{Curl}\,[P\nu_m],$$

so that

$$\frac{dD}{dt} = \frac{1}{4\pi}\frac{dE}{dt} + \frac{dP}{dt} + \text{Curl}\,[P\nu_m],$$

The equation expressing the rate of change of the magnetic force is analogously

$$\frac{dH}{dt} = \frac{dB}{dt} - 4\pi\frac{dI}{dt} - 4\pi\,\text{Curl}\,[I\nu_m].$$

This latter equation must be specially emphasised as it has apparently never yet been explicitly introduced in the theory, although it is necessary to secure greater consistency in the dynamical theory. The expression $\frac{dI}{dt}$ represents the rate of change of the magnetic polarisation at a fixed point in the field only when the magnetic media as a whole are at rest. When these media are in motion there will be a contribution to this rate due to convection just as in the electric case, and the argument for its exact form may be developed on the same lines. The vectors B and H are, so to speak, attached to the æther, just as were the vectors D and E,† whilst the vector I, as the vector P, is attached to the matter and moves with it. The last two equations contain, therefore, an explicit expression of the effect of the motion so that they are in a sense more convenient than the equations

---

* *Cf.* my 'Theory of Electricity,' pp. 365–367, or LARMOR, 'Æther and Matter,' Chap. IV.

† The vector H being the composite vector of the magnetic theory is analogous to the vector D of the electric theory; the æthereal vector B is analogous to the æthereal vector E. This is the reverse of the usual convention, but see below § 10.

$$\mathrm{D} = \frac{1}{4\pi}\mathrm{E}+\mathrm{P}, \qquad \mathrm{H} = \mathrm{B}-4\pi\mathrm{I},$$

defining merely the values of D and H at any point: they are, of course, ultimately consistent with these equations for, taking the second one as an example, we have

$$\frac{d}{dt}\,\mathrm{div}\;\mathrm{H} = \frac{d}{dt}\,\mathrm{div}\;\mathrm{B}-4\pi\frac{d}{dt}\,\mathrm{div}-4\pi\;\mathrm{div}\;\mathrm{Curl}\;[\mathrm{I}\nu_m],$$

or

$$\frac{d}{dt}\,\mathrm{div}\,(\mathrm{H}-\mathrm{B}+4\pi\mathrm{I}) = 0.$$

With the possible exception of the equation defining $d\mathrm{H}/dt$ it is now generally agreed that the scheme here presented provides a completely effective specification of the kinematical connexions in the electromagnetic field.

To obtain some idea of the effect of these connexions on the dynamical processes operative in the field a further assumption is necessary, and this may take one of several forms which will be reviewed in the sequel. For the present we are concerned merely with these equations as effective representatives of the electromagnetic processes. They are sometimes given another form, by the introduction of a scalar potential ϕ and a vector potential A, these being such that

$$\mathrm{B} = \mathrm{Curl}\;\mathrm{A}, \qquad \mathrm{E} = -\frac{1}{c}\frac{d\mathrm{A}}{dt}-\mathrm{grad}\;\phi,$$

with the other two equations

$$\mathrm{Curl}\;\mathrm{H} = \frac{4\pi}{c}\mathrm{C}, \qquad \mathrm{div}\;\mathrm{E} = 4\pi\rho.$$

The first two of these equations are equivalent to the remaining fundamental equation of MAXWELL'S theory which they replace, but they suffer from the serious disadvantage that the quantities A and ϕ specified in them are not completely defined by the equations as given and require additional data to fix them.

4. We have just noticed that an additional assumption of a dynamical character is necessary to render the Maxwellian electromagnetic scheme completely effective as an electrodynamic theory. The simplest and most direct form of this assumption may be taken to be that expressing the force of electrodynamic origin acting on an arbitrarily moving element of charge, this force being, per unit charge equal to

$$\mathrm{F} = \mathrm{E}+\frac{1}{c}\,[\nu\mathrm{B}]$$

v being the velocity of the charge and square brackets being employed as throughout to denote the vector product. This expression, first given in its complete form by LORENTZ and LARMOR, is generally regarded as being exact.*

But there are other forms of the additional assumption which are equally effective and in some respects more general than that just given. We may, for instance, assume definite expressions for the potential and kinetic energies of electromagnetic origin in the system and combine these with the assumption that the dynamical processes operative in the field are governed by the same general laws as are the processes in a similar mechanical system. This form of the argument proves ultimately to be consistent with the first as regards the expression for the effective mechanical force on a moving element of charge, but it has the advantage of being expressed in more general terms, thus carrying with it the possibility of fitting better with any modification that it may subsequently be thought desirable to make in our general conception of the theory. It need not be presumed that this form of the argument is any less general than the first on account of the fact that it apparently involves more than one assumption, for this increase in number is counterbalanced by the fact that the dynamical argument ultimately reduces the two fundamental equations to one, FARADAY'S relation being derived simultaneously as a consequence of AMPÈRE'S.

The general dynamical argument was first formulated by MAXWELL† for the case of the field surrounding a system of linear currents. His analysis was subsequently extended to cover the more general case, firstly by VON HELMHOLTZ‡ and LORENTZ,§ later by LARMOR,|| MACDONALD,¶ ABRAHAM,** and others. In the later investigations the whole subject is regarded from the point of view of the theory of electrons, wherein every manifestation of the field is regarded as arising in the aggregate disposition or motion of electronic charges; even in the former investigations, although they are apparently of a more general character, certain assumptions are involved which render their analyses theoretically effective only under the same restricted circumstances. In each of these cases it is assumed, practically speaking, that the potential energy of the electromagnetic field is represented by the expression

$$W = \frac{1}{8\pi}\int E^2\,dv$$

and the kinetic energy by

$$T = \frac{1}{2c}\int (AC)\,dv,$$

---

* Relatively modifications are not here under consideration.

† 'Treatise,' vol. 2, Ch. VI., VII.

‡ 'Ann. Phys. Chem.,' vol. 47 (1892), p. 1.

§ 'La Théorie Électromagnétique de MAXWELL' (Leiden, 1892), §§ 55–61.

|| 'Æther and Matter,' § 50.

¶ 'Electric Waves,' Appendix I.

** 'Ann. Phys.,' vol. 10 (1903), p. 105.

both volume integrals being taken throughout the entire field: the derivation of the dynamical and field equations is then accomplished by an application of one or the other of the well-known processes of analytical dynamics. The interpretation of the same results for the case when the kinetic energy is given by the usual expression

$$T = \frac{1}{8\pi}\int B^2\,dv$$

was given by the present author.*

Whilst the theoretical simplicity of these discussions, which results from their interpretation in terms of the simple electronic hypothesis, is a great point in their favour, it seemed, nevertheless, of theoretical interest at least to attempt to formulate the problem under less restricted conditions, especially in view of the pronounced tendency exhibited in some quarters to deny the adequacy of the Maxwellian theory as a complete microscopic theory. Besides the more general discussion in the form in which it is here presented emphasises certain difficulties inherent in the usual formulations which have not hitherto received adequate attention.

5. The most general dynamical principle which determines the motion of every material system is the Law of Least Action expressible in the usual form

$$\delta\int_{t_1}^{t_2} L\,dt \equiv \delta\int_{t_1}^{t_2}(T-W)\,dt = 0$$

wherein T denotes the kinetic energy and W the potential energy of the system in any configuration and formulated in terms of any co-ordinates that are sufficient to specify the configuration in accordance with its known properties and connexions, and where the variation refers to a fixed time of passage of the system from the initial to the final configuration. This is the ordinary form of HAMILTON'S principle, but it involves in any case a complete knowledge of the constitution of the systems, because before it can be applied it is necessary to know the exact values of the kinetic and potential energies expressed properly in terms of the co-ordinates and velocities. As however we have frequently to deal with systems whose ultimate constitution is either wholly or partly unknown it is necessary to establish a modified form of the principle allowing for a possible ignorance of the constitution of the systems with which we may have to deal. The modification is fully discussed in most works on analytical dynamics,† and we may here content ourselves by merely presenting the results, interpreting them however in a manner somewhat different from that usually given, in order to throw some light on certain questions which arise in the subsequent application in our present theory. Suppose then that it has been found impracticable to express the Lagrangian function L in terms of the chosen co-ordinates of the system, the typical one of which we may denote by q; but that it is expressed in terms of a certain number of variables $x_1, x_2, \ldots x_k$, which are known to be connected with the co-ordinates q and their velocities $\dot{q}$ by a series of relations of the type

---

* 'Phil. Mag.,' vol. 32 (1916), p. 195.

† *E.g.*, 'Treatise on Analytical Dynamics' (2nd ed., Cambridge, 1918), by E. T. WHITTAKER.

$$M_s = 0$$

M_s being a function of the co-ordinates q, the velocities $\dot{q}$, the variables x and the differential coefficients of these latter variables with respect to the time. For the sake of simplicity we shall restrict our statement to the case when the first differentials only appear. The usual method of procedure is to introduce a set of multipliers λ_s, functions of the time, and then to consider the variations of the integral

$$\int_{t_1}^{t_2} (L+\Sigma\lambda_s M_s)\, dt$$

where the q's and x's undergo independent variations. The equations obtained for the vanishing of the variation are of two types. Firstly, there is an equation of the type

$$\frac{d}{dt}\left(\Sigma_s \lambda_s \frac{\partial M_s}{\partial \dot{x}}\right) - \Sigma\lambda_s \frac{\partial M_s}{\partial x_s} - \frac{\partial L}{\partial x} = 0$$

for each variable x: these with the restricting equations will determine the x's and λ's as functions of the co-ordinates q and the time. Then there is an equation of the type

$$\frac{d}{dt}\left(\Sigma_s \lambda_s \frac{\partial M_s}{\partial \dot{q}}\right) - \Sigma\lambda_s \frac{\partial M_s}{\partial q} = 0$$

for the motion in each q co-ordinate.

The latter equations only involve the Lagrangian function L through the quantities λ and x which enter into it, and once these are determined the rest of the solution involves only the restricting conditions. In fact when once these multipliers and variables are determined and regarded as functions of the time only the motion in the q co-ordinates is completely determined by the condition that the integral

$$\int \Sigma\lambda_s M_s \, dt$$

is stationary for independent variations of the co-ordinates q. It may even happen that the relations M involve the co-ordinates q and the variables x in such a way that it is possible to separate M into two terms, one of which is a function explicitly of the q's only and the other of the x's only. In this case the part of the integral required in the above statement is only that part of it involving the q's and this is independent entirely of the co-ordinates x.

This remark has an important bearing on a question which occurs in the sequel, and it shows that the existence of a variational form for the equations of motion does not necessarily imply that the integrand involved is a true Lagrangian function for the system.

6. Now let us apply these principles to our electromagnetic problem. The conditions in the field surrounding a number of bodies are specified in the usual way by the magnetic induction vector B and the electric force vector E, and the part of the Lagrangian function associated with this field may be taken to be

$$\frac{1}{8\pi}\int (B^2 - E^2)\, dv,$$

the first term denoting the magnetic or kinetic energy and the second the electric or potential, and the integral is extended over the whole of space. In addition to these energies there will be the energies of the material bodies in the field which will consist in part of the kinetic energies of their organised motions, in part of their potential energy relative to one another or to any extraneous fields of non-electric nature, and in part finally of internal energy of elastic or motional type in the media. The part of the Lagrangian function corresponding to these energies can, in the most general case, be denoted by

$$\int (L_0 - W_i)\, dv$$

where L_0 is the Lagrangian function of the organised motions of the media, reckoned per unit volume at each place and assumed to be a function only of the position co-ordinates and velocities, and W_i is the internal energy of all types reckoned as potential energy per unit volume: this latter term will be a function of the electric and magnetic polarisations in the media, but will be assumed not to depend to any appreciable extent on the rates of variation of these conditions, and in so far as some of the internal energy is essentially of kinetic type, it will be in reality a sort of modified Lagrangian function with the energy corresponding to the motional terms converted to potential energy in the usual way. The function L_0 may also be taken to include a part arising from the assumed inertia of any free electrons that may be present.

The motion of the system can now be expressed in the form

$$\delta \int_{t_1}^{t_2} dt \int \left[L_0 - W_i + \frac{1}{8\pi}(B^2 - E^2)\right] dv$$

and we could conduct the variation directly were it not for the fact that our functions are not all expressed explicitly in terms of the independent co-ordinates of the systems, which are in reality the position co-ordinates of the elements of matter and electricity. As indicated above we can however avoid the use of any such explicit interpretation by the use of undetermined multipliers. In this way the variations of E and B can be temporarily rendered independent of each other and of the actual co-ordinates of the material and electrical elements.

We know that the vectors of the theory are connected with one another and the actual co-ordinates of the system by the equations

$$\operatorname{div} \mathrm{E} = 4\pi\Sigma e - 4\pi \operatorname{div} \mathrm{P}$$

$$\operatorname{Curl} \mathrm{H} = \frac{1}{c}\frac{d\mathrm{E}}{dt} + \frac{4\pi}{c}\frac{d\mathrm{P}}{dt} + \frac{4\pi}{c}\operatorname{Curl}\left[\mathrm{P}\dot{r}_m\right] + \frac{4\pi}{c}\Sigma e\left(\dot{r}_m + \dot{r}_e\right)$$

$$\frac{d\mathrm{B}}{dt} = \frac{d\mathrm{H}}{dt} + 4\pi\frac{d\mathrm{I}}{dt} + 4\pi \operatorname{Curl}\left[\mathrm{I}\dot{r}_m\right].$$

In these equations P is the dielectric and I the magnetic polarisation intensity; $\dot{r}_m$ is the velocity and r_m the position vector of the element of matter and $\dot{r}_e$ and r_e the velocity and position vectors relative to this element of the typical element of free charge (e) over which the sum Σ in the first and second equations is taken *per unit volume* at each place.

In these equations we have purposely refrained from assuming a definite electronic constitution for the dielectric and magnetic polarisations as it was desired to emphasise certain points in connexion with the mechanical forcive which have not yet been adequately dealt with.

We have thus to introduce three undetermined multipliers one scalar ϕ and two vectors A_1, A_2 and it is then the variation of

$$\int_{t_1}^{t_2} dt \int dv \left[\mathrm{L}_0 - \mathrm{W}_i + \frac{1}{8\pi}\left(\mathrm{B}^2 - \mathrm{E}^2\right) + \frac{\phi}{4\pi}\left(\operatorname{div} \mathrm{E} + 4\pi \operatorname{div} \mathrm{P}\right)\right.$$
$$- \frac{1}{4\pi}\left(\mathrm{A}_1, \operatorname{Curl} \mathrm{H} - \frac{1}{c}\frac{d\mathrm{E}}{dt} - \frac{4\pi}{c}\frac{d\mathrm{P}}{dt} - \frac{4\pi}{c}\operatorname{Curl}\left[\mathrm{P}\dot{r}_m\right]\right)$$
$$+ \frac{1}{4\pi c}\left(\mathrm{A}_2, \frac{d\mathrm{B}}{dt} - \frac{d\mathrm{H}}{dt} - 4\pi\frac{d\mathrm{I}}{dt} - 4\pi \operatorname{Curl}\left[\mathrm{I}\dot{r}_m\right]\right)$$
$$\left. - \Sigma e\phi + \frac{1}{c}\Sigma e\left(\mathrm{A}_1, \dot{r}_m + \dot{r}_e\right)\right]$$

that is to be made null, afterwards determining the forms of the various undetermined functions to satisfy the restrictions which necessitated their introduction. In conducting the variation we can now treat the electric force, displacement and polarisation, the magnetic force induction and polarisation and the position co-ordinates of the electrical and material elements as all independent. We here see the reason for introducing the equation expressing the rate of change of H instead of the equation

$$\mathrm{H} = \mathrm{B} - 4\pi\mathrm{I}$$

determining its value, for this latter equation does not in reality enable us to obtain a relation between the variations of H and the position co-ordinates of the matter, so we could not treat all our variables as independent.

In the main the details of the variational calculation possess no novel features and need not here be elaborated. There are, however, one or two terms which require careful handling, especially when finding the variations due to the alteration of position of the matter.

The variations of terms of the type

$$\int \phi \operatorname{div} \mathrm{P}\, dv, \quad \frac{1}{c}\int\left(\mathrm{A}_1 \frac{d\mathrm{P}}{dt}\right) dv, \quad \frac{1}{c}\int (\mathrm{A}_1\, \mathrm{Curl}\, [\mathrm{P}\dot{r}_m])\, dv,$$

due to variations in the position co-ordinates of the elements of matter to which the vectors P and $\dot{r}_m$ are attached should not be performed until the differential operators affecting the P function are eliminated by an integration by parts. If we bear this in mind we shall find that the final result for the variation consists of terms at the time and space limits, which require separate adjustment, together with

$$\begin{aligned}\int_{t_1}^{t_2} dt \int dv \Bigg[\delta \mathrm{L}_0 &+ \frac{1}{4\pi}\left(\delta \mathrm{B},\ \mathrm{B} - \frac{1}{c}\frac{d\mathrm{A}_2}{dt}\right) - \frac{1}{4\pi}\left(\delta \mathrm{E},\ \mathrm{E} + \nabla\phi + \frac{1}{c}\frac{d\mathrm{A}_1}{dt}\right) \\ &- \frac{1}{4\pi}\left(\delta \mathrm{H},\ \mathrm{Curl}\ \mathrm{A}_1 - \frac{1}{c}\frac{d\mathrm{A}_2}{dt}\right) - \left(\mathrm{P},\ (\delta r_m \nabla)\left(\nabla\phi + \frac{1}{c}\frac{d\mathrm{A}_1}{dt}\right)\right) \\ &- \frac{1}{c}\left(\delta r_m,\ \left[\mathrm{Curl}\ \mathrm{A}_1,\ \frac{d\mathrm{P}}{dt}\right]\right) - \frac{1}{c}\left(\delta r_m \left[\mathrm{Curl}\ \frac{d\mathrm{A}_1}{dt},\ \mathrm{P}\right]\right) \\ &- \frac{1}{c}(\delta r_m \nabla)(\mathrm{Curl}\ \mathrm{A}_1,\ [\mathrm{P}\dot{r}_m]) + \frac{1}{c}\left(\mathrm{I},\ (\delta r_m \nabla)\frac{d\mathrm{A}_2}{dt}\right) \\ &+ \frac{1}{c}\left(\delta r_m,\ \left[\mathrm{Curl}\ \mathrm{A}_2,\ \frac{d\mathrm{I}}{dt}\right]\right) + \frac{1}{c}\left(\delta r_m \left[\mathrm{Curl}\ \frac{d\mathrm{A}_2}{dt},\ \mathrm{I}\right]\right) \\ &+ \frac{1}{c}(\delta r_m \nabla)(\mathrm{Curl}\ \mathrm{A}_2,\ [\mathrm{I}\dot{r}_m]) \\ &+ \Sigma e\left(\delta r_m + \delta r_e,\ -\nabla\phi - \frac{1}{c}\frac{d\mathrm{A}_1}{dt} + \frac{1}{c}[\dot{r}_m + \dot{r}_e,\ \mathrm{Curl}\ \mathrm{A}_1]\right) \\ + \delta \mathrm{W}_i &+ \left(-\nabla\phi - \frac{1}{c}\frac{d\mathrm{A}_1}{dt} + \frac{1}{c}[\dot{r}_m,\ \mathrm{Curl}\ \mathrm{A}_1],\ \delta \mathrm{P}\right) \\ &\qquad + \left(\frac{1}{c}\frac{d\mathrm{A}_2}{dt} - \frac{1}{c}[\dot{r}_m,\ \mathrm{Curl}\ \mathrm{A}_2],\ \delta \mathrm{I}\right)\Bigg],\end{aligned}$$

where in the terms $(\delta r_m \nabla)$ (Curl A_1, $[\mathrm{P}\dot{r}_m]$) and $(\delta r_m \nabla)$ (Curl A_2, $[\mathrm{I}\dot{r}_m]$) the vector operator ∇ $\left(\text{whose components are } \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ is presumed to affect only the functions A_1 and A_2.

The variations δr_m, δr_e which determine the virtual displacements of the electrical and material elements and the variations δE, δB, δP, …, can now be considered as all independent and perfectly arbitrary, and hence the coefficient of each must vanish separately in the dynamical variational equation. The coefficients of the latter variations lead to the equations

$$B = \frac{1}{c}\frac{dA_2}{dt}, \qquad E = -\nabla\phi - \frac{1}{c}\frac{dA_1}{dt},$$

$$\text{Curl } A_1 = \frac{1}{c}\frac{dA_2}{dt},$$

from which it follows that the multipliers ϕ and A_1 are the ordinary scalar and vector potentials of the theory so that further

$$\text{Curl } E = -\frac{1}{c}\frac{d}{dt}\text{Curl } A_1.$$

As regards the vector A_2, this is a new vector potential whose *curl* is required in our subsequent discussions. For this we have

$$\text{Curl } B = \text{Curl } H + 4\pi\,\text{Curl } I = \frac{1}{c}\frac{d}{dt}\text{Curl } A_2,$$

whilst if we use C_e as the total current of true electric flux we have, by AMPÈRE'S equation

$$\text{Curl } H = \frac{1}{c}\frac{dE}{dt} + \frac{4\pi}{c}C_e.$$

Thus, if we use

$$C'_e = C_e + c\,\text{Curl } I,$$

we have

$$\frac{d}{dt}(\text{Curl } A_2) = \frac{dE}{dt} + 4\pi C'_e.$$

The main part of *Curl* A_2 is therefore represented by the electric force: there is however in addition a local term E_0 depending on the time integral of the current density at the point. We can thus write

$$\text{Curl } A_2 = E + E_0.$$

If we use the values thus determined for the various undetermined multipliers introduced at the outset, the remaining terms of the variation give for the motion of the material and electrical elements equations of the type

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial L_0}{\partial \dot{x}_m}\right) - \frac{\partial L}{\partial x_m} = {} & \left(P\frac{\partial E}{\partial x_m}\right) + \left(I\frac{\partial B}{\partial x_m}\right) + \frac{1}{c}\left([P\dot{r}_m],\ \frac{\partial B}{\partial x_m}\right) \\ & - \frac{1}{c}\left([I\dot{r}_m],\ \frac{\partial(E+E_0)}{\partial x_m}\right) - [P,\ \text{Curl } E]_x \\ & - [I,\ \text{Curl } B]_x - \frac{1}{c}\left[B\frac{dP}{dt}\right]_x + \frac{1}{c}\left[E+E_0,\ \frac{dI}{dt}\right]_x \\ & + \rho\left(E_x + \frac{1}{c}[\dot{r}_m B]_x\right) + \frac{1}{c}[C_1 B]_x,\end{aligned}$$

for the motion of the matter and of type

$$\frac{d}{dt}\left(\frac{\partial L_0}{\partial \dot{x}_m}\right) - \frac{\partial L}{\partial x_m} = e\left(E_x + \frac{1}{c}[\nu B]_x\right)$$

for the electrical elements. In the first of these equations $\rho = \Sigma e$ is the density of the free charge, and $C_1 = \Sigma e\dot{r}_e$ is the density of the true conduction current; in the second equation $r = r_m + r_e$ is the absolute position co-ordinate of the electrical element and $\nu = \dot{r}$ is its resultant velocity.

In addition we have the variational equation for the internal energy which can be left as it stands in the form

$$\delta W_i = \left(E + \frac{1}{c}[\dot{r}_m B],\ \delta P\right) + \left(B - \frac{1}{c}[\dot{r}_m,\ E + E_0],\ \delta I\right).$$

Interpreted in terms of the language of ordinary dynamics these equations imply that

$$(P\nabla)\,E_x + (I\nabla)\,B_x + \frac{1}{c}\left([P\dot{r}_m],\ \frac{\partial B}{\partial x}\right) - \frac{1}{c}\left([I\dot{r}_m],\ \frac{\partial(E+E_0)}{\partial x}\right)$$
$$+\frac{1}{c}\left[\frac{dP}{dt} + C_1,\ B\right]_x - \frac{1}{c}\left[\frac{dI}{dt},\ E + E_0\right]_x + \rho\left(E_x + \frac{1}{c}[\dot{r}_m B]_x\right)$$

is the force per unit volume at each place tending to accelerate the motion of the matter, whilst

$$e\left(E + \frac{1}{c}[\nu B]\right)$$

is the force tending to accelerate the motion of the element of electricity e.

The electrical terms in the first of these results are consistent with those obtained by LARMOR* and others, but the magnetic terms, which are in fact analogous with the electrical terms depending on the dielectric polarisation, are fundamentally different from those obtained by these authors. Further, since

$$E = -\frac{1}{c}\frac{dA_1}{dt} - \nabla\phi,$$

it follows that

$$\text{Curl } E = -\frac{1}{c}\frac{d}{dt}\text{Curl } A_1$$
$$= -\frac{1}{c}\frac{dB}{dt},$$

so that our equations are in complete agreement with the most general form of MAXWELL's theory.

---

* 'Æther and Matter,' Chap. VI.

7. If we examine the above analyses closely we shall notice a rather important point bearing on a fundamental question which has already been the subject of some discussion.* If we take the integral in its complete form with the variation carried out and with the values of the various multipliers inserted it can be seen to reduce to the expression

$$\begin{aligned}\int_{t_1}^{t_2} dt \int dv \Big[& \delta \mathrm{L}_0 - \delta \mathrm{W}_i + (\mathrm{P}\,\delta \mathrm{E}) + (\mathrm{I}\,\delta \mathrm{B}) - \frac{1}{c}((\delta \dot{r}_m \nabla),\ \mathrm{B},\ [\mathrm{P}\dot{r}_m]) - (\delta \dot{r}_m,\ [\mathrm{P}\ \mathrm{Curl}\ \mathrm{E}]) \\ & - \frac{1}{c}((\delta r_m \nabla)(\mathrm{E}+\mathrm{E}_0),\ [\mathrm{I}\dot{r}_m]) - (\delta r_m [\mathrm{I}\ \mathrm{Curl}\ \mathrm{B}]) \\ & + \left(\frac{\delta \dot{r}_m}{c},\ \left[\frac{d\mathrm{P}}{dt},\ \mathrm{B}\right]\right) - \frac{1}{c}\left(\delta \dot{r}_m,\ \left[\frac{d\mathrm{I}}{dt},\ \mathrm{E}+\mathrm{E}_0\right]\right) \\ & + \left(\mathrm{E} + \frac{1}{c}[\dot{r}_m \mathrm{B}],\ \delta \mathrm{P}\right) + \left(\mathrm{B} - \frac{1}{c}[\dot{r}_m,\ \mathrm{E}+\mathrm{E}_0],\ \delta \mathrm{I}\right) \\ & + \Sigma e\left(\delta \dot{r}_m + \delta \dot{r}_e,\ \mathrm{E} + \frac{1}{c}[\mathrm{VB}]\right)\Big],\end{aligned}$$

in which neither the electric nor magnetic energy contributes an explicit term. This is the first definite indication we have that the modified function with which we may operate to find the equations of motion of the electric and material elements is explicitly independent of the expression for the energy in the æthereal field. We may, in fact, see that, just as in the dynamical problem examined above, the whole circumstances of the motion in the real co-ordinates of the system can be derived by the variational principle, using the integral

$$\begin{aligned}\int_{t_1}^{t_2} dt \int dv \Big[& \mathrm{L}_0 - \mathrm{W}_i - \phi\ \mathrm{div}\ \mathrm{P} - \frac{1}{c}\left(\mathrm{A}_1 \frac{d\mathrm{P}}{dt}\right) - \frac{1}{c}(\mathrm{A}_1\ \mathrm{Curl}\ [\mathrm{P}\dot{r}_m]) \\ & - \frac{1}{c}\left(\mathrm{A}_2 \frac{d\mathrm{I}}{dt}\right) - \frac{1}{c}(\mathrm{A}_2\ \mathrm{Curl}\ [\mathrm{I}\dot{r}_m]) + \Sigma e\phi + \Sigma e(\mathrm{A}_1,\ \dot{r}_m + \dot{r}_e)\Big]\end{aligned}$$

just as we used the Hamiltonian integral, taking in it E, A_1, A_2, ϕ as functions of the time and space co-ordinates only. It is of course possible to establish this directly, for it is easily verified that the difference between the integrand just employed and the previous one, viz.,

$$\mathrm{L}_0 - \mathrm{W}_i + \frac{1}{8\pi}(\mathrm{B}^2 - \mathrm{E}^2),$$

involves only complete differentials with respect to the time or space co-ordinates. This difference therefore integrates out to the limits and remains ineffective as regards the general dynamical variational equations, and we can therefore use either integrand indiscriminately.

---

* *Cf.* 'Phil. Mag.,' vol. 32 (1916), p. 195, where references to previous work are given.

This is the general form of the result obtained by SCHWARZSCHILD* that in the special case when we are concerned only with free electrons moving in an æthereal field free from matter their equations of motion can be derived by the variational principle, using the integral

$$\int_{t_1}^{t_2} dt \left[\mathrm{L}_0 - \Sigma e\phi + \Sigma \frac{1}{c} (\mathrm{A}\dot{r}) \right],$$

where ϕ, A, the ordinary scalar and vector potentials, are regarded as functions of the time and space variables only.

We now see why it is that consistent formulæ have been obtained by different authors using apparently different expressions to represent the field energies. The results are in fact all explicitly independent of any particular interpretations for these energy expressions.

8. The general dynamical formulation of § 6 agrees with the fact that the material media of the field have an internal constitution which enables them to resist the setting up of the electric and magnetic polarisations by forces $\mathrm{E} + \frac{1}{c}[\dot{r}_m \mathrm{B}]$ and $\mathrm{B} - \frac{1}{c}[\dot{r}_m, \mathrm{E} + \mathrm{E}_0]$ respectively, and that in consequence of any change in the polarised state of these media their intrinsic energy of elastic or motional type is increased by the amount

$$\int dv \int_1^2 (\mathrm{E}' \, \delta\mathrm{P}) + (\mathrm{B}' \, \delta\mathrm{I}),$$

where we have used

$$\mathrm{E}' = \mathrm{E} + \frac{1}{c}[\dot{r}_m \mathrm{B}], \qquad \mathrm{B}' = \mathrm{B} - \frac{1}{c}[\dot{r}_m, \ \mathrm{E} + \mathrm{E}_0].$$

Thus in setting up the electric field and its associated dielectric polarisations in the medium the potential energy of the field is increased by the amount

$$\frac{1}{8\pi} \int \mathrm{E}^2 \, dv.$$

on account of the establishment of the æthereal field together with

$$\int dv \int (\mathrm{E}' \, \delta\mathrm{P})$$

on account of the material polarisation, both amounts being reckoned as potential energy.

This gives a total for the field equal to

$$\int dv \int \left\{ \frac{1}{4\pi} (\mathrm{E} \, \delta\mathrm{E}) + (\mathrm{E} \, \delta\mathrm{P}) + \frac{1}{c} ([\dot{r}_m \mathrm{B}] \, \delta\mathrm{P}) \right\}.$$

---

* 'Gött. Nachr. (Math.-phys. Kl.),' 1903, p. 125.

This result is consistent with that generally obtained in these theories. The last term arising on account of the current due to the convection of the polarisations is however probably of kinetic origin.

Of course, in the general case, all the potential energy put into the field cannot be got out of it again in the form of useful mechanical work, or, in other words, it is not all available. The effectively available energy in the present case consists in the main of the part

$$\int dv \left(\frac{\mathrm{E}^2}{8\pi} + \int (\mathrm{P}\, \delta\mathrm{E}) \right).$$

For the magnetic polarisations the results are somewhat different. In this case the kinetic energy of the field is assumed to be

$$\frac{1}{8\pi} \int \mathrm{B}^2\, dv,$$

to which we must add the intrinsic kinetic energy involved in the induced magnetic polarisations to obtain the total energy in the field; reckoned as potential energy the intrinsic energy is

$$\int dv \int (\mathrm{B}'\, d\mathrm{I}),$$

and therefore as kinetic energy it is

$$-\int dv \int (\mathrm{B}'\, d\mathrm{I}),$$

giving a total for the field equal to

$$\frac{1}{8\pi} \int \mathrm{B}^2\, dv - \int dv\, (\mathrm{B}'\, d\mathrm{I})$$
$$= \frac{1}{4\pi} \int dv \int (\mathrm{B}\, d\mathrm{H}) + \frac{1}{c} \int dv \int (\mathrm{E} + \mathrm{E}_0,\ [d\mathrm{I},\ \dot{r}_m]),$$

a result which is again practically equivalent to that usually given in this theory.

If we treat the convection of the dielectric polarisation as effectively equivalent to a magnetic polarisation of intensity

$$\mathrm{I}' = \frac{1}{c} [\mathrm{P} \dot{r}_m],$$

and the convection of the magnetic polarisation as effectively equivalent to a dielectric polarisation of intensity

$$\mathrm{P}' = -\frac{1}{c} [\mathrm{I} \dot{r}_m],$$

and include the energies corresponding to these two parts in with the appropriate totals the formulæ just obtained are somewhat simpler in form. The total kinetic energy will be now

$$\frac{1}{8\pi}\int\left[\mathrm{B}^2-8\pi\int(\mathrm{B}\,d\mathrm{I})-8\pi\int(\mathrm{B}\,d\mathrm{I}')\right]dv$$

or

$$\int\frac{dv}{4\pi}\int(\mathrm{B}\,d\mathrm{H}'),$$

where

$$\mathrm{B} = \mathrm{H}'+4\pi\mathrm{I}+\frac{4\pi}{c}[\mathrm{P}\nu_m],$$

and the potential energy is

$$\frac{1}{8\pi}\int\left[\mathrm{E}^2+8\pi\int(\mathrm{E}\,\delta\mathrm{P})+8\pi\int(\mathrm{E}\,\delta\mathrm{P}')\right]dv,$$

or is

$$\iint(\mathrm{E}\,d\mathrm{D}')\,dv,$$

where

$$\mathrm{D}' = \frac{\mathrm{E}}{4\pi}+\mathrm{P}+\mathrm{P}'.$$

This, however, leaves out of account the term

$$\int dv\int(\mathrm{E}_0\,\delta\mathrm{P}'),$$

which, representing as it does the energy of local actions and reactions, may be assumed to be inoperative as regards general mechanical processes, and may in consequence be rejected altogether.

In connexion with the discussions of the expressions for the internal energy of the material media it may be worth while emphasising the significance of the various terms in the expression for the forcive on the polarised material elements derived in the previous paragraph, especially the way in which the various terms arise and their dependence on the term

$$\mathrm{Curl}\,[\mathrm{P}\nu_m]$$

in the complete current. Further, we may notice how the complete expression for the change of intrinsic energy per unit volume, viz.,

$$\left(\mathrm{E}+\frac{1}{c}[\nu_m\,\mathrm{B}],\,\delta\mathrm{P}\right)$$

is derived, the latter term being due to the above-mentioned term in the current.

This last remark points to the possibility of obtaining an elementary deduction of the expression

$$\mathrm{E}+\frac{1}{c}[\nu_m, \mathrm{B}]$$

for the complete electromotive forcive simply by calculating the rate of change of intrinsic energy of a moving bi-pole, and the calculation has in fact been carried out by LARMOR,* taking however a parallel plate condenser with equal and oppositely charged plates, moving in a uniform magnetic field.

An analogous argument in the magnetic case will give a deduction of the magnetomotive forcive

$$\mathrm{B}-\frac{1}{c}[\nu_m, \mathrm{E}].$$

9. We have stated that the magnetic energy expressions just obtained are effectively equivalent to those usually derived, whereas as a matter of fact this is true only of the final result; the various formulæ employed in the derivation of this result are not in their usual form but it has been shown elsewhere† that they are consistent with the complete dynamical theory, the more usual formulæ and the various modifications of them which have from time to time been suggested being all inadmissible on this score. A complete discussion of the justification for this last statement is necessarily beyond the scope of the present paper, but it may perhaps serve a useful purpose if a brief outline of some of the more important reasons is given, especially as they have some bearing on points raised elsewhere in the present discussion.

In the first place there is probably little or no difficulty in seeing the fallacy in the usual and simplest form of the theory wherein the expression $\mu\mathrm{H}^2/8\pi$ for the magnetic energy density is derived in the statical theory as potential energy and in the dynamical theory as kinetic energy: we need only enquire as to the type of energy represented by the same expression when the field is due in part to rigid magnets and in part to steady currents. The more consistent result is obtained by taking

$$\frac{\mu\mathrm{H}^2}{8\pi}-\frac{1}{4\pi}(\mathrm{HB})$$

as the expression for the density in the statical case as this agrees with the opposite sign in the dynamical case and yet gives the same total.

There is however another form of the results first tentatively suggested by HERTZ and HEAVISIDE and subsequently developed in great detail by other writers, more particularly by R. GANS‡ and H. WEBER, wherein the difficulty presents itself in another and more involved form. The fundamental basis of this theory is the assumption of a distribution of true magnetic matter of density at any place equal to

---

* 'Proc. Lond. Math. Soc.' (1915).

† *Cf.* my 'Theory of Electricity,' p. 417, or 'Roy. Soc. Proc.,' vol. 93, A, p. 20 (1916).

‡ 'Ann. der Physik,' vol. 13 (1904), p. 634, and 'Encyklopädie der Math. Wissensch.,' vol. v., art. 15, where references to other authors will be found.

$$\rho_m = -\operatorname{div} \mathrm{I}_0 = \frac{1}{4\pi} \operatorname{div} \mathrm{B}$$

wherein I_0 is the density of the permanent magnetic polarity. This magnetic matter is supposed to be distributed continuously throughout the space but so that the amount in any portion of the matter is zero, a condition which is perhaps rather difficult of realisation, as it would make the distribution in any particular portion of the matter dependent on the distribution in all the surrounding portions.

In this theory the magnetic energy is first calculated on analogy with the electrostatic energy; the magnetic induction vector B is regarded as a sort of composite displacement produced by the acting force H, so that the energy per unit volume is

$$\frac{1}{4\pi}\int^{\mathrm{B}} (\mathrm{H}\, d\mathrm{B}).$$

This expression is then verified to be equivalent in the purely statical case to the volume integral

$$\int dv \int^{\rho_m} \phi\, d\rho_m$$

taken over the entire field, the surface integral over the infinite boundary contributing nothing in all regular cases; ϕ is the magnetic potential of the field.

In generalising the theory to the case where the field is due to linear currents the same physical basis is adopted as regards the expression

$$\int dv \int^{\mathrm{B}} (\mathrm{H}\, d\mathrm{B}),$$

which still therefore remains valid, and when there are no permanent magnets about this is easily verified by the usual argument to be equivalent to the summation

$$\frac{1}{c}\Sigma \int^{\mathrm{N}} \mathrm{J}\, d\mathrm{N}$$

over the different current elements, J denoting the typical current strength and N the induction through its circuit. When there are permanent magnets present this expression becomes

$$\int dv \int^{\rho_m} \phi\, d\rho_m + \frac{1}{c}\Sigma \int^{\mathrm{N}} \mathrm{J}\, d\mathrm{N}.$$

It is then shown that the mechanical forcive on the magnetic matter in any one of its co-ordinates is derivable as the appropriate *negative* gradient of this energy function, whilst the force on a current is to be obtained as the *positive* gradient with respect to its position co-ordinate.

Unfortunately all the authors concerned merely talk of magnetic energy without specifying whether it is to be taken as kinetic or potential energy. One might perhaps infer that as the results are interpreted in terms of a static potential function it is implied that all the energies are potential, but the fact that the forces on the currents are derivable as the positive gradients of the function

$$\frac{1}{c}\Sigma\int^{N} J\,dN$$

suggests that this part of the energy at least is kinetic energy. The difficulty of sign is therefore still present.

Even if we confine ourselves to the statical theory the same interpretation is not entirely free from difficulties of another kind. The potential energy in the field is taken to be represented by

$$\int dv\int^{\rho_m}\phi\,d\rho_m,$$

but this expression really represents the total energy in the field; in the general case the only part of this energy which is mechanically available is

$$\int dv\int^{\phi}\rho_m\,d\phi,$$

and this is properly speaking the potential function from which the mechanical forces acting on the magnetism are to be derived. Of course, when the law of induction is linear the intrinsic energy of the field is equal to the available energy, but even then their natures are fundamentally different and equality in their magnitudes is hardly a sufficient justification for confusing the one with the other.

Apart from this difficulty, however, the next step employed in the development of the theory will cause some trouble. To effect the transformation from the expression

$$\int dv\int^{\rho_m}\phi\,d\rho_m$$

to the equivalent expression

$$-\int dv\int^{I_0}(H\,dI_0)$$

the method of integration by parts is employed. But LARMOR has shown that two expressions of this type being derived the one from the other by the method of integration by parts, really represent fundamentally different distributions of the energy in the field, although the total amounts represented by them are the same. The two expressions cannot therefore be used indiscriminately to determine the stresses around an element of the magnetic matter. It is not, of course, possible at the present stage to say which of the two expressions does represent the true energy distribution, but any examination of the mechanical inter-relations of the different magnetic masses in the field postulates a previous decision as to the proper expression to be taken as representing the available energy of these masses in its normal form, respecting both its total amount and proper distribution. Once this decision has been made it is unsafe to employ the method of integration by parts unless due regard is paid to the surface integrals thereby introduced.

The distribution of energy interpreted in terms of ideal magnetic matter which is properly equivalent to the expression

$$-\int dv \int^{I_0} (H\, dI_0)$$

is such that the energy in any volume of the field consists of a distribution throughout it of density at each point equal to

$$-\int \phi \operatorname{div} dI_0 = -\int \phi\, d\rho_m$$

together with a surface distribution of density

$$\int \phi\, dI_n$$

over its surface. This corresponds properly to POISSON'S distribution of magnetic matter and emphasises the necessity for the inclusion of surface distributions of magnetic matter.

This explains why it is that the above theory determines properly the forces on the permanent magnets as a whole, but fails to give a consistent account of the internal forces between different parts of the same magnet. At the surface of an ordinary magnet it may quite legitimately be assumed that owing to the existence of a transition layer, the normal component of the magnetisation vanishes there, and consequently the surface integrals applied to the outer surfaces of any such body would also vanish; the two different expressions for the contained energy thus become equivalent.

This is perhaps sufficient to justify a summary rejection of this new interpretation of the energy relations of the magnetic field, as being at most no better than the older one which it presumed to displace. The real trouble in both cases seems to have arisen mainly in an effort to discover an analogy in the relations of the electric and magnetic fields. HERTZ and HEAVISIDE were the first to insist on the existence of this analogy, and practically all the modern writers follow them in this matter, even so far as to regard it as providing sufficient justification for certain fundamental formulæ of the theory. A close scrutiny of the subject will, however, reveal the fact that although the mathematical relations connecting the magnetic force induction and polarisation are to a certain extent similar to those connecting the electric force, total electric displacement and dielectric displacement, the similarity ends with these relations, and the dynamical characteristics of the two types of field are essentially different ; and the analogy itself, so far as it exists, seems to be based on erroneous and confused conception of the nature of the magnetic energy as determined by the usual expression of the theory, so that it finds no place in a consistent formulation of the subject, notwithstanding even HEAVISIDE's spirited defence in criticism of LARMOR.

In his treatise ABRAHAM adopts the analogy as a sufficient basis for the discussion of the magnetic theory, but decides that the procedure is not without its difficulties, particularly as regards the ferro-magnetic phenomena ; not being able to overcome these he condemns the whole procedure as being inadequate to include a proper account of these matters.

10. We now turn from these discussions to a brief review of the general energy relations of the electromagnetic field. A concise account of these relations so far, that is, as they have been dealt with in existing accounts of the subject, has been given with full references in my treatise (Ch. XIV.), and it will suffice for the present to give the barest outlines of the discussion so far as they may be required.

The fundamental equation expressing in its most concise form the energy principle for the electromagnetic field can be written in the form

$$\frac{dW}{dt}+\frac{dT}{dt}+F+\int S_n\,df=0.$$

wherein W and T are respectively the potential and kinetic energies inside any volume bounded by the closed surface f in the field, F is the rate of dissipation of electromagnetic energy into energy of other types such, for instance, as results mainly from the inertia of the electric charges constituting the conduction and convection currents ; and S is the vector determining the flow of electromagnetic energy outwards over the bounding surface.

In this equation it is generally assumed that our knowledge of the forms of F and W is precise and accurate, and that in fact in agreement with the results of paragraph 8.

$$W=\frac{1}{8\pi}\int E^2\,dv+\int dv\int^{P}(E'\,\delta P)$$

where P is the polarisation intensity of the dielectric media produced by the electromotive force

$$E'=E+\frac{1}{c}[V_m B]$$

whilst

$$F=\int(EC_1)\,dv$$

where C_1 represents the part of the total current depending on the motion of the electrons constituting the conduction current and the current due to the convection of electric charges, but excluding the part due to the convection of the polarised media.

It seems difficult to dispute the form of the expression for W, but careful consideration will also convince one that it is probably just as difficult to support it in the most general case, except it be by the results which are derived from it, which certainly seem to be in satisfactory agreement with our knowledge of these things. A similar reservation must be applied also to the expression for F, but there is here an additional point worth noticing. It is not often remarked that the form given tacitly involves an assumption which is derived as an independent result from discussions based on this special form. In fact it involves the definite assumption that no work is done by the magnetic forces during the motion of electric charges. Of course the usual expression for such force as proportional to the vector product of the velocity and magnetic force confirms this assumption, but the derivation of this expression by dynamical methods from results derived from the present discussion is by so much deprived of interest. In fact, if to the assumption that these forces do no work we add the further conditions that they are linear in the magnetic and velocity vectors, it would appear that their form is completely determined, at least to a constant factor, without further considerations either of a dynamical or any other nature.

The form for the expression T is not usually regarded as being sufficiently definite to be used in the present connexion, mainly because it is the more readily convertible into equally simple alternative forms. We have in our previous discussions made certain assumptions which have proved to be equivalent to taking

$$\mathrm{T} = \frac{1}{4\pi}\int dv \int^{\mathrm{B}} (\mathrm{H}\, d\mathrm{B}).$$

but this special form will subsequently be proved to be irrelevant to the discussion. It is usually regarded as being most advisable to consider the expression for T as bound up with that for S, the equation connecting the two being that derived from the energy principle by the insertion of the forms chosen for W and F, viz.,

$$\frac{d\mathrm{T}}{dt} + \int_f \mathrm{S}_n\, df = -\int (\mathrm{EC})\, dv - \int \frac{1}{c}\left([v_m \mathrm{B}],\ \frac{d\mathrm{P}}{dt} \right) dv + \frac{1}{c}\int (\mathrm{E},\ \mathrm{Curl}\, [\mathrm{P} v_m])\, dv$$

C being now the total current of MAXWELL's theory. This is all we can derive from the energy principle. The various possibilities open to us have been examined in detail before. We may take

$$\mathrm{Curl}\, \mathrm{H} = \frac{4\pi}{c}\, \mathrm{C}$$

and then we get POYNTING'S theory in which

$$T = \frac{1}{4\pi}\int dv\left[\int^{B}\left(H-\frac{4\pi}{c}[P\nu_m],\ dB\right)-\frac{1}{c}\int^{P}\left([\nu_m B]\ dP\right)\right]$$

and

$$S = \frac{c}{4\pi}\left[E,\ H-4\pi[P\nu_m]\right] = \frac{c}{4\pi}[E,\ H']$$

where

$$H' = H-4\pi[P\nu_m]$$

is equivalent to the vector H′ introduced in paragraph 8.

This theory has the advantage, in addition to that already discussed at length, that it involves no further dynamical assumption other than those expressed in the special forms chosen for W and F. The Amperean equation used to effect the separation being more in the nature of a kinematical definition of the electric current or magnetic force than of a dynamical relation between the field vectors.

Another form can be obtained by using the equations

$$E = -\frac{1}{c}\frac{dA}{dt} - \text{grad}\,\phi$$

with

$$\text{div}\ C = 0$$

we then get

$$T = \frac{1}{c}\int dv\int^{A}(C\,dA)-\frac{1}{c}\int dv\int^{P}(\nu_m[B\,dP])$$

$$-c\int^{A}(\text{Curl}\,[P\nu_m].\ dA)$$

with

$$S = \phi\,(C-c\,\text{Curl}\,[P\nu_m]).$$

The special form of this result when the media are at rest has been shown* to be inconsistent with our usual conception of such things as radiation phenomena.

Yet another form of the theory can be obtained by taking

$$T = -\int dt\int\left[(EC+\frac{1}{c}\left([\nu_m B].\ \frac{dP}{dt}\right)-c\,(E\ \text{Curl}\,[P\nu_m])\right]dv$$

and then

$$S = 0.$$

In such a theory there would be no such thing as radiation.

We can go on multiplying the different forms of this theory indefinitely and each form obtained would in itself be perfectly consistent with the Maxwellian electrodynamic theory. The expressions for S and T in them are of course dependent on each other, being connected by the equation given above, and there will be a relation between the forms for two different theories. In fact if S and T are the forms corresponding to any one mode of separation and if we write

---

* 'Phil. Mag.,' vol. 34 (1917), p. 385.

$$S' = S + \frac{dU}{dt}$$

where U is any arbitrary vector function we shall have

$$\frac{dT}{dt} + \int S_n \, df = \frac{dT}{dt} + \int \left(S'_n - \frac{dU_n}{dt} \right) df$$

$$= \frac{dT'}{dt} + \int S'_n \, df$$

where

$$T' = T - \int \operatorname{div} U dv,$$

and S′ and T′ are appropriate forms for a new mode of separation. In this way, by assigning convenient values for U, we might tentatively construct a number of interesting formulæ.

The last result also shows why it is that the particular form chosen for the kinetic energy is irrelevant to the general dynamical discussion of paragraph 7. In fact, if, instead of the form T used on that occasion, we had employed the general value derived above

$$T - \int \operatorname{div} U \, dv$$

the part of the variation depending on this energy becomes the time integral of

$$\delta T - \int \operatorname{div} \delta U \, dv,$$

and the latter integral reduces to a surface integral over the infinitely distant boundary and cannot therefore contribute anything in this general variational equation.

Of course, from another point of view, the various forms of the theory here under review, differ merely in assigning different distributions to the magnetic energy in the field, each of these distributions being ultimately consistent with the same proper total for this quantity; and the fact that they all lead to the same dynamical equations, merely verifies a well-known result of analytical dynamics that the particular form of expression for the energies of the system is immaterial to the ultimate dynamical equations for the field inside a continuous medium. Of course the solutions of boundary problems such as are, for example, involved in a specification of the energy flux, depend essentially on the particular form assumed for the energy distribution ;

and it has been shown on a previous occasion* that the only form of specification of the energy distribution which is consistent with our usual ideas on these matters is that which makes the density of the magnetic energy equal to $B^2/8\pi$, and as the present discussion shows that our only hope of discrimination lies in that direction, we may assume that the evidence in favour of this special form is conclusive, at least for the present; it is besides the only form in which the most general case is completely representative of the distribution in any volume of the field without requiring the introduction of boundary terms involving surface distributions.

11. We next turn to a consideration of the expression for the forcive of electromagnetic origin acting on the polarised media in the field. We have seen that the mechanical forcive on the dielectrically polarised media is such that its x-component per unit volume at any place is of the form

$$(\mathrm{P}\nabla)\,\mathrm{E}_x+\frac{1}{c}\left([\mathrm{P}\nu_m],\,\frac{\partial \mathrm{B}}{\partial x}\right)+\frac{1}{c}\left[\frac{d\mathrm{P}}{dt},\,\mathrm{B}\right]_x$$

or as it first appears in the analysis

$$\left(\mathrm{P}\frac{\partial \mathrm{E}}{\partial x}\right)-[\mathrm{P}\,\mathrm{Curl}\,\mathrm{E}]_x+\frac{1}{c}\left[\frac{d\mathrm{P}}{dt},\,\mathrm{B}\right]+\frac{1}{c}\left([\mathrm{P}\nu_m]\,\frac{\partial \mathrm{B}}{\partial x}\right)$$

$$=\mathrm{P}\frac{\partial \mathrm{E}}{\partial x}+\frac{1}{c}\frac{d}{dt}[\mathrm{PB}]+\frac{1}{c}\left([\mathrm{P}\nu_m]\,\frac{\partial \mathrm{B}}{\partial x}\right).$$

This result is in complete agreement with that derived by LARMOR in the electron theory,† but the present derivation indicates clearly the origin of the different terms in it. The expression

$$\left(\mathrm{P}\frac{\partial \mathrm{E}}{\partial x}\right)$$

is that corresponding to the expression derived in the statical theory from energy considerations and corresponds to MAXWELL'S magnetic expression; the second term, viz.,

$$-[\mathrm{P}\,\mathrm{Curl}\,\mathrm{E}]_x$$

is one of the terms arising as a result of the convection of the media, and this is the term which is effective in reducing the electric part of the forcive to the form

$$(\mathrm{P}\nabla)\,\mathrm{E}_x$$

which is the result derived in the elementary theory by regarding the forcive as the resultant of the forces on the elementary bi-poles.

---

* 'Phil. Mag.,' vol. 34 (1917), p. 385. *Cf.* also 'Phil. Mag.,' vol. 32 (1916), p. 162.

† 'Æther and Matter,' p. 104.

Similar results apply in the magnetic case. The general expression for the forcive on the magnetic media is, per unit volume, equal to

$$\left(\mathrm{I}\frac{\partial \mathrm{B}}{\partial x}\right)-[\mathrm{I}\ \mathrm{Curl}\ \mathrm{B}]_x-\frac{1}{c}\left([\mathrm{I}\nu_m],\ \frac{\partial(\mathrm{E}+\mathrm{E}_0)}{\partial x}\right)$$

$$-\frac{1}{c}\left[\frac{d\mathrm{I}}{dt},\ \mathrm{E}+\mathrm{E}_0\right]_x$$

$$=(\mathrm{I}\nabla)\,\mathrm{B}_x-\frac{1}{c}\left([\mathrm{I}\nu_m]\frac{\partial(\mathrm{E}+\mathrm{E}_0)}{\partial x}\right)-\frac{1}{c}\left[\frac{d\mathrm{I}}{dt},\ \mathrm{E}+\mathrm{E}_0\right]$$

$$=\left(\mathrm{I}\frac{\partial \mathrm{B}}{\partial x}\right)-\frac{1}{c}\frac{d}{dt}[\mathrm{I},\ \mathrm{E}+\mathrm{E}_0]-\frac{1}{c}\left([\mathrm{I}\nu_m],\ \frac{\partial(\mathrm{E}+\mathrm{E}_0)}{\partial x}\right).$$

In these expressions the parts depending on E_0 and I^2, representing as they do forces on the elements of the media determined solely by the conditions in those elements, would be neglected in a mechanical theory.* The expression for the effective forcive thus reduces to

$$\left(\mathrm{I}\frac{\partial \mathrm{H}}{\partial x}\right)-\frac{1}{c}\frac{d}{dt}[\mathrm{IE}]-\frac{1}{c}\left([\mathrm{I}\nu_m],\ \frac{\partial \mathrm{E}}{\partial x}\right)$$

$$=(\mathrm{I}\nabla)\,\mathrm{H}-\frac{1}{c}\left([\mathrm{I}\nu_m]\frac{\partial \mathrm{E}}{\partial x}\right)-\frac{1}{c}\left[\frac{d\mathrm{I}}{dt},\ \mathrm{E}\right].$$

This expression is only equivalent to MAXWELL's expression in the statical case he considers. It is, however, practically equivalent to that derived by counting the forces on the constituent poles, but even here the general result rather suggests a modified conception of the force on a magnetic pole, this force involving in the general case a term due to the electric force. The question of the existence of forces on a magnetic pole due to its motion in an electric field does not appear to have been investigated on an independent basis, although it is definitely contained in the relations of transformation involved in the theory of relativity, which require the form for this forcive

$$\mathrm{B}-\frac{1}{c}[\nu_m\mathrm{E}].$$

It will however be proved below in the next paragraph that such forces do probably exist and are in fact of precisely the correct type.

It may, of course, be objected that the last term in the equation

$$\frac{d\mathrm{B}}{dt}=\frac{d\mathrm{H}}{dt}+4\pi\frac{d\mathrm{I}}{dt}+4\pi\ \mathrm{Curl}\ [\mathrm{I}\nu_m],$$

which is the origin of the discrepancy obtained for the magnetic forcive, does not in reality exist, but yet the other results derived from this equation are almost certainly indisputable, and it seems difficult to realise a state of affairs where the equation without the last term would be generally true.

---

* *Cf.* LARMOR, 'Æther and Matter,' p. 98.

The whole question of mechanical forces on polarised media is ultimately bound up with the question of the variation of the intrinsic energy of those media, and the expression

$$\left(\mathrm{I}\frac{\partial \mathrm{H}}{\partial x}\right),$$

for the x-component of the forcive per unit volume implies that the internal energy of those media change in a small displacement by

$$(\mathrm{H}\,\delta\mathrm{I})$$

per unit volume. But when the media are in motion the expression for the change $\delta\mathrm{I}$ used in this expression involves a part due to the convection of the polarisation which is more properly concerned with the mechanical forces than with the intrinsic elastic or motional ones, as it would exist if the internal constitution of the media was maintained rigidly constant. It is therefore suggested that the result derived above, that the expression for the rate of change of the intrinsic energy is practically

$$\left(\mathrm{H}-\frac{1}{c}[\nu_m\mathrm{E}].\,\delta\mathrm{I}\right)$$

per unit volume, is the more legitimate form of this expression, as allowance is made in it for the convection, and if this is granted, then the equivalent expression for the mechanical forcive, viz.,

$$(\mathrm{I}\nabla)\,\mathrm{H}_x-\frac{1}{c}\left([\mathrm{I}\nu_m]\frac{\partial \mathrm{E}}{\partial x}\right)-\frac{1}{c}\left[\frac{d\mathrm{I}}{dt}\,.\,\mathrm{E}\right]$$

must be regarded as the only adequate form.

Moreover these two expressions essentially involve the particular form of equation adopted for defining $d\mathrm{H}/dt$, and are the only ones which are capable of fitting in with a general relativity theory.

The results here derived also emphasise the difficulties involved in treating the currents due to the convection of polarised media as effectively equivalent to a polarisation of the opposite kind. If, for example, we had treated the convection current

$$\mathrm{Curl}\,[\mathrm{P}\nu_m]$$

as equivalent to a distribution of magnetic polarity of intensity

$$[\mathrm{P}\nu_m]$$

at each place from the outset we should have been led to an entirely erroneous expression for the forcive on the polarised media, the reason being that the inclusion of this quantity with the magnetism hides its true character, particularly as regards its dependence on the velocity of the medium.

Nevertheless many of the relations of the theory will be considerably simplified if this procedure is adopted.

12. The question of forces on fictitious magnetic poles moving in an electric field is easily resolved by imparting to such poles a substantiality sufficient to allow us to talk of forces on them, and then applying any of the general methods used in this theory. The Lagrangian function of the system is still of the form

$$\mathrm{L}+\frac{1}{8\pi}\int(\mathrm{B}^2-\mathrm{E}^2)\,dv,$$

L being that part of it which is not directly determined by the conditions in the field and which will as usual be assumed to be a function of the positions and velocities of the electric and magnetic elements only. The sequence of changes is then best described by the fact that the action integral

$$\int_{t_1}^{t_2} dt\left[\mathrm{L}+\frac{1}{8\pi}\int(\mathrm{B}^2-\mathrm{E}^2)\,dv\right]$$

taken between fixed time limits is stationary subject to the implied conditions of the field. If we assume generally that there are a number of discrete electric particles as well in the field, these conditions may be written in the form

$$\mathrm{div}\ \mathrm{E}-4\pi\Sigma e = 0,$$

$$\mathrm{div}\ (\mathrm{B}-\mathrm{H})+4\pi\Sigma m = 0,$$

$$\mathrm{Curl}\ \mathrm{H}-\frac{1}{c}\frac{d\mathrm{E}}{dt}-\frac{4\pi}{c}\Sigma e v_e = 0,$$

$$\frac{d\mathrm{B}}{dt}-\frac{d\mathrm{H}}{dt}-4\pi\Sigma m v_m = 0,$$

wherein e is the charge of the typical electron and v_e its velocity, m is the strength of the typical magnetic particle whose velocity is v_m and the sum Σ in each equation is taken per unit volume at each place over the respective elements indicated in it.

We now introduce four undetermined Lagrangian multiplying functions, two scalar quantities ϕ_1 and ϕ_2, and two vector quantities A_1 and A_2, it is thus the variation of

$$\int dt\left[\mathrm{L}+\frac{1}{8\pi}\int dv\left[\mathrm{B}^2-\mathrm{E}^2+2\phi_1\,\mathrm{div}\ \mathrm{E}+2\phi_2\,\mathrm{div}\ (\mathrm{B}-\mathrm{H})\right.\right.$$

$$\left.-2\left(\mathrm{A}_1,\mathrm{Curl}\ \mathrm{H}-\frac{1}{c}\frac{d\mathrm{E}}{dt}\right)+\frac{2}{c}\left(\mathrm{A}_2\frac{d\mathrm{B}}{dt}-\frac{d\mathrm{H}}{dt}\right)\right]$$

$$\left.-8\pi\Sigma e\phi_1+8\pi\Sigma m\phi_2+\frac{8\pi}{c}\Sigma e\,(\mathrm{A}_1 v_e)-\Sigma m\,(\mathrm{A}_2 v_m)\right],$$

that is to be made null, afterwards determining the functions ϕ_1, ϕ_2, A_1, A_2 to satisfy the restrictions which necessitated their introduction. The variation can now be affected in the usual way, and the condition that it vanishes leads to the following equations

$$E + \text{grad } \phi_1 + \frac{1}{c}\frac{dA_1}{dt} = 0,$$

$$B - \text{grad } \phi_2 - \frac{1}{c}\frac{dA_2}{dt} = 0,$$

with

$$\text{Curl } A_1 - \text{grad } \phi_2 - \frac{1}{c}\frac{dA_2}{dt} = 0,$$

with three equations of each of the types

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_e}\right) - \frac{\partial L}{\partial x_e} = -e\left(\frac{\partial \phi_1}{\partial x_e} + \frac{1}{c}\frac{dA_{1x}}{dt}\right) + \frac{e}{c}[v_e \text{ Curl } A_1].$$

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}_m}\right) - \frac{\partial L}{\partial x_m} = m\left(\frac{\partial \phi_2}{\partial x_m} + \frac{1}{c}\frac{dA_{2x}}{dt}\right) - \frac{m}{c}[v_m \text{ Curl } A_2].$$

The first and third of these equations show that ϕ_1 and A_1 are the usual scalar and vector potentials; in fact from the third we have

$$\text{Curl } A_1 = \text{grad } \phi_2 + \frac{1}{c}\frac{dA_2}{dt}$$
$$= B,$$

so that A_1 is the vector potential and then ϕ_1 is the scalar potential.

The fourth equation thus determines the usual expression for the reaction forces on the moving electron; the fifth equation determines similarly the force on the moving magnetic pole in the form

$$m\left(\text{grad } \phi_2 + \frac{1}{c}\frac{dA_2}{dt}\right) - \frac{m}{c}[v_m \text{ Curl } A_2] = m\left(B - \frac{1}{c}[v_m \text{ Curl } A_2]\right).$$

We have yet to determine Curl A_2: we have

$$B = \text{grad } \phi_2 + \frac{1}{c}\frac{dA_2}{dt}.$$

so that

$$\text{Curl } B = \frac{1}{c}\frac{d}{dt}\text{ Curl } A_2.$$

Now

$$\text{Curl } B = \frac{4\pi}{c}C_1 + 4\pi \text{ Curl } I + \frac{1}{c}\frac{dE}{dt} = \frac{4\pi}{c}C_0 = \frac{1}{c}\frac{dE}{dt},$$

wherein C_1 is the true current density of electric flux ;

$$I = \frac{B-H}{4\pi},$$

is the intensity of magnetic polarity, and

$$C_0 \equiv C_1 + c \text{ Curl } I.$$

It follows that

$$\text{Curl } A_2 = E + 4\pi \int^t C_0 \, dt = E + E_0,$$

say. The vector A_2 is a slightly more general form of the second vector potential introduced in our previous dynamical discussion and its *curl* is identical with the *curl* of that vector.

The main part of Curl A_2 is thus determined by the electric force in the field, and its mechanically effective part is completely represented by this vector ; the forcive on the moving magnetic pole is thus to all intents and purposes equal to

$$m\left(B - \frac{1}{c}\,[\nu_m E]\right)$$

an expression which agrees with that suggested by the relativity transformation.

It must, however, be noted that the local term is necessary in the complete relation defining the vector A_2 for the simpler relation

$$\text{Curl } A_2 = E,$$

carries with it the consequence that

$$\text{div } E = 0$$

at all points of the field, and this is true only of those points where there is no electricity.

The expression for the forcive on the magnetic media is now attainable by regarding it as the resultant of the forces on its contained poles ; for the volume v bounded by the closed surface f, it is in fact

$$-\int (\text{div } I)\, B \, dv + \int I_m B \, df$$

$$+ \frac{1}{c}\int \left[\frac{dI}{dt} + c \text{ Curl } [I\nu_m], E + E_0\right] dv$$

$$+ \int [\,[\,[I\nu_m]\, n_1], E + E_0]\, df.$$

The second and fourth integrals transform by GREEN's lemma to the volume integrals

$$\int (B\,(\nabla I) + (I\nabla)\,B)\, dv - \int [\,[\text{Curl } [I\nu_m]\, E + E_0] + \text{grad } ([I\nu_m]\, E + E_0)$$

$$+ \qquad\qquad -([I\nu_m] \text{ div } (E + E_0)]\, dv,$$

where in the last term but one the gradient operation only affects the E functions. Now

$$\text{div}\,(\text{E}+\text{E}_0) = 0,$$

and thus the resultant force may be taken as distributed throughout the volume with intensity at each point equal to

$$(\text{I}\nabla)\,\text{B}+\frac{1}{c}\left[\frac{d\text{I}}{dt},\ \text{E}+\text{E}_0\right]-\text{grad}\,(\,[\,\text{I}v_m],\ \text{E}+\text{E}_0)$$

in agreement with the general result derived above. The local terms in I^2 and E_0 may again be presumed to balance out with other forces of a type not at present under review.

13. The two new potentials A_2 and ϕ_2, introduced in the analysis of the last paragraph, are the general forms of the potentials analogous to the ordinary scalar and vector potentials of this theory, and they satisfy similar equations. We have, in fact,

$$\text{Curl A}_2 = \text{E}+4\pi\int^t \text{C}_0\,dt$$

where C_0 is the total current density of electric flux including the effective representation of the magnetism. Thus

$$\text{Curl Curl A}_2 = \text{Curl E}+4\pi\int^t \text{Curl C}_0\,dt$$

$$= -\frac{1}{c}\frac{d\text{B}}{dt}+4\pi\int^t \text{Curl C}_0\,dt.$$

Thus

$$\text{grad div A}_2-\nabla^2\text{A}_2 = -\frac{1}{c^2}\frac{d^2\text{A}_2}{dt^2}-\frac{1}{c}\,\text{grad}\,\frac{d\phi_2}{dt}+4\pi\int^t \text{Curl C}_0\,dt$$

whilst since $\text{div B} = 0$, we have also

$$\nabla^2\phi_2+\frac{1}{c}\frac{d}{dt}\,\text{div A}_2 = 0.$$

We may now adopt one of a number of alternatives. The simplest one is got by taking $\phi_2 = 0$, when we also have

$$\text{div A}_2 = 0$$

with, therefore,

$$\nabla^2\text{A}_2 = \frac{1}{c^2}\frac{d^2\text{A}_2}{dt^2}+4\pi\int^t \text{Curl C}_0\,dt.$$

The last equation really involves the first, for

$$\nabla^2\,(\text{div A}_2) = \frac{1}{c^2}\frac{d^2}{dt^2}(\text{div A}_2).$$

so that div A_2 must be zero as it has no singularities.

The vector A_2 chosen in this way is, practically speaking, the æthereal displacement vector employed by **Larmor** in his mechanical model of the electric and luminiferous medium. The *curl* of this vector is the electric force, or at least as regards its rate of change, whilst the magnetic induction B, which is proportional to the time rate of change of A_2, appears as the velocity.

We need not, however, take the quantities in this way. We might take

$$\text{div } A_2 = \frac{1}{c}\frac{d\phi_2}{dt} - 4\pi \text{ div} \int^t \mathrm{I}\, dt,$$

and then we should have

$$\nabla^2\phi_2 = \frac{1}{c^2}\frac{d^2\phi_2}{dt} - \text{div I}$$

with

$$\nabla^2 A'_2 = \frac{1}{c^2}\frac{d^2 A'_2}{dt^2} + 4\pi\frac{d\mathrm{I}}{dt} + 4\pi\int^t \text{Curl } C_0\, dt$$

where we have used

$$A'_2 = A_2 - 4\pi\int^t \mathrm{I}\, dt.$$

In this case ϕ_2 is the scalar potential of the magnetic distribution. whilst A'_2 belongs to the current distribution. With these differential equations the general values of ϕ_2 and A_2 in regular fields are such that

$$\phi_2 = \frac{4\pi}{c}\int \frac{[\text{div I}]}{r}\, dv$$

whilst

$$\frac{1}{c}\frac{dA_2}{dt} - \frac{4\pi \mathrm{I}}{c} = \frac{4\pi}{c^2}\int r^{-1}\left[\text{Curl } C_0 + \frac{d^2\mathrm{I}}{dt^2}\right] dv$$

the square brackets in the integrands denoting that their values are taken at each point for the time $\left(t-\frac{r}{c}\right)$.

These are the most interesting cases of the solutions for ϕ_2 and A_2, but we may construct any number of others. It must be noticed, however, that the equation

$$B = \frac{1}{c}\frac{dA_2}{dt} + \text{grad } \phi_2,$$

does not imply that the vector B is derived from a potential in steady fields, for it is impossible to satisfy the equations with A_2 independent of the time; we may have dA_2/dt constant in time but not A_2. This is the origin of the difficulty in **Larmor's** mechanical model which seems to necessitate the piling up of æthereal displacement in a steady magnetic field.

14. We have determined the complete expression for the forcive per unit volume on the media occupying the electromagnetic field. The next step in the general

theory is to reduce these forces to a representation by means of an applied stress system of ordinary character. This discussion leads in the usual way to the introduction of the concept of electromagnetic momentum.

The actual calculations for the present form of the results are not materially different from those given *in extenso* elsewhere, so that it will again be sufficient to outline the principal stages in the discussion. The method employed is to attempt to express, say, the x component of the force per unit volume in the form

$$\frac{\partial T_{xx}}{\partial x}+\frac{\partial T_{xy}}{\partial y}+\frac{\partial T_{xz}}{\partial z}.$$

Now the total forcive of electrodynamic origin acting on the medium of the field at any place is such that its x component per unit volume is

$$\begin{aligned}&\left(P-\frac{1}{c}[I\dot{r}_m], \frac{\partial E}{\partial x}\right)+\left(I+\frac{1}{c}[P\dot{r}_m], \frac{\partial B}{\partial x}\right)\\ &\quad-[P, \text{Curl } E]_x-[I, \text{Curl } B]_x-\frac{1}{c}\left[B.\frac{dP}{dt}\right]+\frac{1}{c}\left[E.\frac{dI}{dt}\right]\\ &\quad+\rho\left(E_x+\frac{1}{c}[\nu_m B]\right)+\frac{1}{c}[C_1 B]_x.\end{aligned}$$

Again writing

$$H' = H-\frac{4\pi}{c}[P\nu_m],$$

it is proved just as in the usual form of the theory that the forcive of which this is the representative component is represented in the main by a stress system in which

$$T_{xx} = E_x D_x-\frac{1}{8\pi}E^2+\frac{1}{4\pi}B_x H'_x-\frac{1}{8\pi}H'^2$$

and

$$T_{xy} E_x D_y+\frac{1}{4\pi}B_{xy} H'_x$$

with symmetrical expressions for the other constituents; but with this representation there is an outstanding part of the complete forcive, viz.,

$$-\frac{1}{4c}\frac{d}{dt}[EB]+\frac{1}{c}\frac{d}{dt}[EI]-\frac{1}{c}\left([I\nu_m].\frac{\partial E}{\partial x}\right) = -\frac{1}{4\pi c}\frac{d}{dt}[EH]-\frac{1}{c}\left([I\nu_m], \frac{\partial E}{\partial x}\right)$$

which cannot be so reduced. The first term in this outstanding part, being a complete differential with respect to the time, is usually taken to represent a part of the complete forcive arising as the kinetic reaction to a rate of change of momentum, and this is the origin of the concept of electromagnetic momentum. This idea is however partly destroyed by the remaining term in the above expression which cannot be developed either as a forcive of ordinary type or as a kinetic reaction to a rate of change of momentum, so that we are rather forced to regard these outstanding terms as pointing to the failure of the ideas from which we set out. This conclusion does not, of course, invalidate the results derived in the simpler electron theory, as the concept of momentum will remain under the simplest conditions as a convenient mathematical expression for the actual result, whatever be its ultimate physical basis.

The present formulation possesses another disadvantage which is apparently not inherent in the simplest presentations of the momentum idea. In the electron theory, as usually developed, the momentum remains as a fundamental quantity and is distributed over the field with the density

$$\frac{1}{4\pi c}[\mathrm{EB}]$$

at each point; this gives it a purely æthereal constitution as the vectors E and B are those which define the conditions in the æthereal field. In the present formulation the vector B is replaced by the vector H which is essentially an auxiliary mechanical vector in the theory; the fundamental nature of the momentum vector is therefore entirely lost. We can, of course, assume that some of the momentum is in reality attached to the matter, and such an assumption has certain points in its favour. The force of electromagnetic origin on the dielectric media for example has an x component which per unit volume is

$$\left(\mathrm{P}\frac{\partial \mathrm{E}}{\partial x}\right)+\frac{1}{c}\left([\mathrm{P}_{v_m}],\frac{\partial \mathrm{B}}{\partial x}\right)-[\mathrm{P}\ \mathrm{Curl}\ \mathrm{E}]_x+\frac{1}{c}\left[\frac{d\mathrm{P}}{dt},\mathrm{B}\right],$$

and this may be written in the form

$$\left(\mathrm{P}\frac{\partial \mathrm{E}}{\partial x}\right)+\frac{1}{c}\left([\mathrm{P}_{v_m}],\frac{\partial \mathrm{B}}{\partial x}\right)+\frac{1}{c}\frac{d}{dt}[\mathrm{PB}].$$

The first two terms appear as those appropriate to the energy function in the statical theory which would be

$$-(\mathrm{PE})-\frac{1}{c}([\mathrm{P}_{v_m}].\ \mathrm{B}),$$

so that the third might be regarded as a kinetic reaction to a rate of change of momentum, which would be distributed throughout the medium with a density

$$\frac{1}{c}[\mathrm{PB}]$$

at each point.

A similar analysis and analogous results hold for the magnetic media.

There is, too, a relation satisfied by the momentum vector which appears in the simplest form of the theory and to which a fundamental significance is attached by some authors, but which is not satisfied by the results of our present discussion. The vector determining the flux of electromagnetic energy has been seen to be

$$S = \frac{c}{4\pi}[E, H']$$

and that determining the momentum is

$$M = \frac{1}{4\pi c}[E, H].$$

In the absence of magnetic media and convective dielectric polarisations these two expressions satisfy the equation

$$M = \frac{1}{c^2}S.$$

but under the most general circumstances this relation is not satisfied.

We have so far conducted the discussions as though the quantity derived as a momentum is unique and definite, whereas, as a matter of fact, this is far from being the case. We saw that the idea of the momentum arose from certain outstanding terms which remained when attempting to reduce the electromotive forces to a representation by a stress system. Now we can give a number of different forms to this reduction and each one carries with it a different expression for the electromagnetic momentum. We can, for instance, write

$$\begin{aligned}\frac{1}{4\pi c}[EH] &= -\frac{1}{4\pi c^2}\left[\frac{dA}{dt}, H\right] - \frac{1}{4\pi c}[\nabla\phi, H] \\ &= -\frac{1}{4\pi c^2}\left[\frac{dA}{dt}, H\right] - \frac{1}{4\pi c}|\nabla\phi, H| + \frac{1}{c_2}\phi C\end{aligned}$$

and the second term in this expression when differentiated with respect to the time might be included in the stress specification. This would leave a new expression for the electromagnetic momentum which is

$$\frac{1}{c_2}\left(\phi C - \frac{1}{4\pi}\left[\frac{dA}{dt}, H\right]\right)$$

a form which would probably be suitable for use in connexion with a theory in which the radiation phenomena are represented by MACDONALD'S form of the theory.

This is not the only alternative to the usual theory for we can construct similarly any number of others. It appears, however, that the usual presentation is probably the simplest possible one, and this is a great advantage in its favour; but subsequent developments of the theory may require a modification, and then it is as well to remember that there are other forms of the theory perfectly consistent with the general relations of the electromagnetic field, both as regards its general and dynamical aspects.

It is hoped in a future communication to examine in detail some of these alternative expressions for the momentum, but so far the results obtained are not of sufficient interest to warrant their discussion at the present stage.

15. It may now be convenient to summarise the conclusions and results of our discussion. The differential theory of **Maxwell** as expressed in the usual way by the equations

$$\text{Curl } \mathrm{H} = \frac{4\pi}{c}\left(\frac{d\mathrm{D}}{dt} + \mathrm{C}_1 - \rho\nu\right) \qquad \text{Curl } \mathrm{E} = -\frac{1}{c}\frac{d\mathrm{B}}{dt} \qquad \text{div } \mathrm{D} = \rho$$

is supplemented by the addition of an equation expressing the rate of change of the magnetic force

$$\frac{d\mathrm{H}}{dt} = \frac{d\mathrm{B}}{dt} - 4\pi\frac{d\mathrm{I}}{dt} - 4\pi \text{ Curl } [\mathrm{I}\nu_m]$$

this equation being analogous to that expressing the rate of change of electric displacement

$$\frac{d\mathrm{D}}{dt} = \frac{1}{4\pi}\frac{d\mathrm{E}}{dt} + \frac{d\mathrm{P}}{dt} + \text{Curl } [\mathrm{P}\nu_m].$$

The fundamental dynamical equations are then derived by a variational principle equivalent to the principle of Least Action in dynamics; in this discussion the assumption of a definite electronic constitution for the dielectric and magnetic polarisations is specially avoided in order to bring out certain points of the theory which have not previously received adequate treatment. In this way, in addition to deriving the equation

$$\text{Curl } \mathrm{E} = -\frac{1}{c}\frac{d\mathrm{B}}{dt},$$

it can be proved that the forces on the media occupying the field consist of several distinct parts. There is firstly a part

$$\rho\left(\mathrm{F} + \frac{1}{c}[\nu_m \mathrm{B}]\right)$$

due to the free charges associated with the typical point of the matter and a part

$$\frac{1}{c}[\mathrm{C}, \mathrm{B}]$$

due to the true conduction current. Due to the dielectric polarisations there is a part whose x component is

$$(\mathrm{P}\nabla)\,\mathrm{E}_x + \frac{1}{c}\left[\frac{d\mathrm{P}}{dt}, \mathrm{B}\right]_x - \frac{1}{c}\left([\mathrm{P}\nu_m]\frac{\partial \mathrm{B}}{\partial x}\right),$$

and the magnetic polarisations give rise to the analogous terms

$$(\mathrm{I}\nabla)\,\mathrm{B}_r - \frac{1}{c}\left[\frac{d\mathrm{I}}{dt},\ \mathrm{E}\right] - \frac{1}{c}\left([\mathrm{I}\nu_m],\ \frac{\partial \mathrm{E}}{\partial x}\right).$$

With the exception of the magnetic terms these results are in general agreement with those usually derived on the basis of the electronic theory, and the discrepancy in the magnetic terms is proved to arise from the inadequacy of the treatment of the magnetic relations in that theory, no allowance being made in it for the convection of the magnetic polarisations.

The results derived from the dynamical theory are then examined in connexion with the usual developments of the theory in regard to radiation phenomena, to the energetic relations of the magnetic media and, finally, to the fundamental problem of the representation of the forces in the field, as an applied stress system and the subsidiary question of electromagnetic momentum. In regard to each of these results are derived which do not differ materially from those usually given, but the slight discrepancies in each case, although probably of little or no practical significance, prove ultimately to be of theoretical importance as helping to justify the fundamental equations on which they are based. The auxiliary conception of electromagnetic momentum is not however completely attained under the most general conditions, although it will still remain to enable us to obtain an effective mode of expressing certain results of the simpler theory; it is probably present in no other capacity in former interpretations of the theory so that this is hardly a disadvantage of the present formulation.

The present theoretical relations require, of course, to be supplemented by the usual empirical laws for the induction of the two polarisations and the conduction current. We have however specially refrained from introducing these relations as it was desired to emphasise the fact that the theory in its complete form is entirely independent of these laws, so that for example it necessarily covers the most complex fields, involving ferromagnetic inductions and polarisations. If we interpret the theory as determining the electrodynamic changes in the system during its transition from one configuration to another even the presence of hysteritic qualities in the inductions will not vitiate its validity. This is, of course, no special advantage attaching to the present form of the theory as it is in reality fundamentally inherent in every interpretation, although it may be hidden by the particular form of expression adopted.

VIII. *The Influence of Molecular Constitution and Temperature on Magnetic Susceptibility.*

Part IV.—Further Applications of the Molecular Field.

By Major A. E. Oxley, *R.A.F.*, *M.A.*, *D.Sc.*, *Mackinnon Student of the Royal Society.*

Communicated by Prof. J. W. Nicholson; *F.R.S.*

Received May 1,—Read June 26, 1919.

Contents.

Page
(1) Introduction 247
(2) On the energy and ultimate tensile strength associated with crystalline media or gels . . 250
(3) A comparison of the elastic constants of diamagnetic, paramagnetic, and ferro-magnetic crystalline media 257
(4) On the change of volume on crystallization interpreted as a magneto-striction effect of the molecular field 259
(5) Further discussion of the nature of the molecular field 267
(6) On a magnetic theory of chemical combination 277
(7) On some anomalies of the Faraday effect 282
(8) Summary of conclusions 286

(1) Introduction.

The present work is a continuation of that published in 'Phil. Trans., Royal Society,' vol. 214, pp. 109–146, 1914 (Parts I. and II.) and vol. 215, pp. 79–103, 1915 (Part III.).* It will lead to clearness in the following development of the subject if a brief résumé of these papers is given. At the same time, I wish to discuss one or two points in connexion with the views which have been previously advanced and the relation between mechanical and molecular theory.

In Part I. the experimental evidence brought forward has justified the hypothesis of molecular distortion enunciated at the outset. We have thereby been led to regard the molecular configuration of a material medium as a distorted one, and this applies particularly to a substance which is crystalline. The extent of this distortion is small, but is sufficient to account for the observed change of specific susceptibility which occurs on crystallization. Such change will naturally depend upon the particular crystalline symmetry assumed by the substance.

The theoretical treatment given in Part II. is an attempt to account for the phenomena observed by extending the electron theory developed by Langevin so

---

* For brevity, reference to these researches is given under Parts I., II., and III.

[Published, April 9, 1920.

as to include the effects of mutual molecular influences. Such a representation of the facts has led to the recognition of a large molecular forcive, in all crystalline media, depending upon the nature and proximity of the molecules in any particular crystalline grouping. The existence of this intrinsic molecular field can only be inferred indirectly. Although the actual properties of the crystalline state depend upon the operation of this field, yet, except in the case of substances of a ferro-magnetic nature, there is no direct experimental evidence which discloses its very great magnitude.

From a theoretical point of view, there seems to be no doubt that the mutual actions of the molecules are represented by enormous internal forcives in all crystalline media. The usual method of determining the forcive at an internal point of the material medium is to take a cavity whose dimensions are small in comparison with ordinary lengths (*e.g.*, 1 cm.) and yet large compared with molecular dimensions. A convenient designation of the dimensions of the cavity is contained in the phrase "physically small."* In molecular theory, the subdivision of the medium into elements is not valid beyond the limits of physical smallness and only in media which are absolutely continuous may the elements be pushed to limits of "mathematical smallness." In a continuous medium our mathematical functions give us an accurate estimate of the forcives and potentials operating at internal points; in a medium composed of discrete particles these same functions give us only an approximate estimate. A discussion of the nearness of the approximation which can be obtained for material media is of great importance from the point of view of our subject. For the liquid state the question has been considered by Sir JOSEPH LARMOR† who found that the part contributed to the forcive at any internal point by the molecules immediately surrounding that point was, on account of rapid motions and irregular distributions of the axes of the molecules, negligibly small. To quote from LARMOR‡: "The general conclusion may be expressed, in an adaptation of CAUCHY's terminology, by the principle that whenever the integrals in the formulæ for mechanical forces on a material medium cease to be convergent, their principal values must be substituted," and again in the footnote to p. 265,

"This statement (*i.e.*, the above quotation) may be considered to be the mathematical expression of the principle of the mutual compensation of molecular forcives, for which, *cf.*, 'Phil. Trans.', A, 1897, p. 260. The principal value of CAUCHY, as regards the completely defined analytical integrals of Pure Mathematics, would be the value at the centre of a minute spherical cavity. But the quantities which, to avoid periphrasis, have been here called integrals, are really summations of contributions from finite though very small, and complexly constituted, polarised molecules; the distribution of these molecules that occupy our minute cavity is entirely unknown and may be continually changing, so that the only possible principal value is the one that omits the contribution of neighbouring molecules altogether."

---

* LEATHEM, "Volume and Surface Integrals used in Physics," 'Cambridge Monographs,' No. 1, p. 5.

† 'Æther and Matter,' p. 261.

‡ *Loc. cit.*, p. 265.

When now we come to the case of crystalline media, the molecules which were removed from our small spherical cavity would affect considerably the value of the forcive at the centre, for there is in this case no averaging out on account of random motions and orientations of the molecules originally occupying the cavity. As we do not know the relative dispositions of the molecules composing the space lattice, or the law of force which is operative between the molecules, it is quite impossible to calculate the value of the intrinsic forcive for a point inside a crystalline medium, and as the method of averages, the application of which is quite satisfactory when the medium is in a fluid state, breaks down in the case of a crystalline medium, it is clear that our only way of progress lies in indirect deduction from experimental facts which record the change of physical properties accompanying the transition from the liquid to the crystalline state. This is the method which has been adopted in the previous portions of this work and it has been shown that the internal forcive at a point of a crystalline medium is extremely large and comparable, if interpreted magnetically, with the molecular field in ferro-magnetic substances (10^7 gauss).

In Part III. it was shown that the potential energy of the molecules forming a crystalline structure was sufficiently large to account for the magnitude of the thermal energy required for fusion. On p. 201 of his "Æther and Matter," LARMOR states: "These various actions (referring to the disturbance of configuration of steady orbits in molecules by action of an applied magnetic field) involve energy terms for each individual molecule, and the sum for all the molecules, if it could be formed, would represent the total energy of the disturbance of the medium. But such a mere aggregate of terms would be of no use for applications to matter in bulk; what we are concerned with there is the mechanical part of the energy, which must be an analytical function of the specification of matter by volume, determined as to mathematical form by the character of the molecular actions, but with coefficients whose values are to be obtained only by direct experiment."

Although for a fluid medium the total energy of the disturbance of the medium, due to the application of a magnetic field, has little significance, yet, in the transition from the liquid to the crystalline state, during which the molecular field becomes operative, the sum total of the energy disturbance of the medium due to the action of this molecular field is representative of the latent thermal energy which is absorbed when the crystalline medium is fused, and has a definite value for each particular substance.

From the point of view of fluids, the intrinsic forcives mutually compensate and the mathematical functions may be treated as analytic, their principal values being taken. In a crystalline structure, however, the functions cannot be treated as analytic. Indirectly we have obtained a measure of the intrinsic forcive in this case with the aid of experimental data. The most we can obtain by direct experiment, however, is a measure of the mean molecular field, which expresses mathematically how the susceptibility of matter *in bulk* depends upon a transition from the liquid to the crystalline state. When thus considering diamagnetic matter in bulk, the large local forcive which has been shown to bind the molecules of the crystalline structure together need not be considered, since for matter *in bulk* its effects are cut out by the mutual compensation of molecular forcives. It is only when we enquire into the *molecular* structure of the crystalline medium, or to changes in this structure, that we pass to the inner limit where the principles of LARMOR and CAUCHY for the fluid state no longer apply. LARMOR remarks : "The result of the integration still however gave us a valid estimate of the effect of the material system *as a whole*, when we bore in mind that the infinite or rather undetermined term entering at the inner limit really represents the part of the result which depends *solely* upon the local molecular configuration ; a part whose actual magnitude could be determined only when that configuration is exactly assigned or known" (*loc. cit.*, p. 125).

It is with this "infinite or rather undetermined term . . . . . . which depends *solely* upon the local molecular configuration" that these researches are mainly concerned. It has been called the local molecular field of the crystalline medium (Part III., p. 83).

(2) ON THE ENERGY AND ULTIMATE TENSILE STRENGTH ASSOCIATED WITH CRYSTALLINE MEDIA OR GELS.

The large intrinsic potential energy associated with a crystalline medium has been discussed in para. 5, Part III., pp. 90–95. It now remains for us to examine the accompanying stresses to see how far the elastic properties of material media may be interpreted in terms of these intrinsic forcives. Consider first the case of a liquid which is gradually cooled in liquid air so that it passes into a glass-hard transparent gel when it arrives at the temperature of the liquid air.

It has been suggested (Part III., p. 81) that the appearance of rigidity in the gel is due to an interlocking of the irregularly shaped molecules (arranged at random) whose thermal agitation is sufficiently reduced. On account of this random orientation of the interlocked molecules the gel will be isotropic. At such a low temperature, also, the molecular motions will be highly constrained so that a particular molecule will present practically the same aspect to the surrounding molecules over a long period. If this is the case, then the local molecular forcive between this and a neighbouring molecule will act in a definite direction and will not be rapidly changing its direction as would be the case with the same molecules at a considerably higher temperature (in the ordinary liquid state). It is clear therefore that between the molecules of the gel at low temperature we shall have a large local forcive in operation, due to the interaction of the magnetic systems or revolving electrons within each molecule, but the direction of the action of this forcive between any pair of molecules will be one of random distribution, as we pass from pair to pair of molecules, although at any given point it is fixed in direction.

On further cooling, the molecules continue to readjust themselves and the rigidity increases until a glass is formed. Still lower temperatures, accompanied by further molecular readjustment due to reduction of amplitude, may result in such a distortion of the lines of force binding the pairs of molecules together, that a completely new pairing of molecules takes place, resulting in spontaneous crystallization accompanied by thermal evolution as the more stable crystalline state is formed.

Whatever may be the nature of the forces which hold the molecules of a liquid together, we have, in addition to these forces, the intrinsic field referred to above, when the substance passes into a rigid gel or crystallizes, and it is due to this intrinsic field that the two latter media show rigidity.

If H_c be this intrinsic field,* I the local intensity of magnetization,† the potential energy per unit volume associated with the gel or crystalline medium will be

$$\tfrac{1}{2}.H_c.I \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

and this will be over and above any potential energy which the molecules of the liquid possess. This is also a measure of the mechanical stress which binds the molecules of the gel or crystalline medium together and determines their rigidity.

In Part III. we have given reasons for locating the source of the local molecular field within the molecule and we found that in the immediate neighbourhood of a molecule the value of this field, as determined from the properties of crystalline media, is of the order 10^7 gauss. However the molecule is orientated, provided that orientation is not variable with time, the local forcive will be of this order of intensity and in some direction determined by the orientations of the two molecules between which it acts. In a gel, as we pass from molecule to molecule, the direction of this stress will be continually changing (fig. 1). Throughout a crystal, on the other hand, its direction will be constant and will in fact be one of the determining factors of a particular form of crystalline symmetry (fig. 1A).

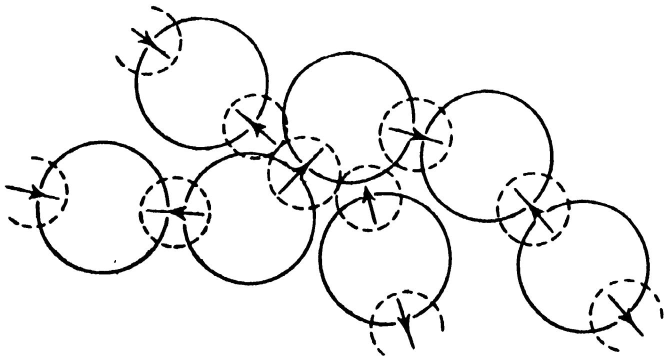


Fig. 1.

In a gel, the whole collection of molecules is bound together into one homogeneous

---

[*] See Part III., p. 86.

[†] *Loc. cit.*, p. 90.

isotropic mass (*i.e.*, as viewed in bulk) whereas in a crystalline medium the mass will be æolotropic.

In Part III. evidence was brought forward showing that in the case of diamagnetic media the local intensity of magnetization I is of the order 100, while the local molecular field H_c between the molecules is of the order 10^7 gauss. The energy of the molecular configuration of the crystalline medium (over and above that due to the molecular configuration in the liquid state) or of the gel at low temperatures, will therefore be of the order $\frac{1}{2} . H_c . I = 10^9$ ergs per unit volume* and the internal stress 10^9 dynes per square centimetre, or 1000 atmospheres approximately. For wrought iron the energy per unit volume will be

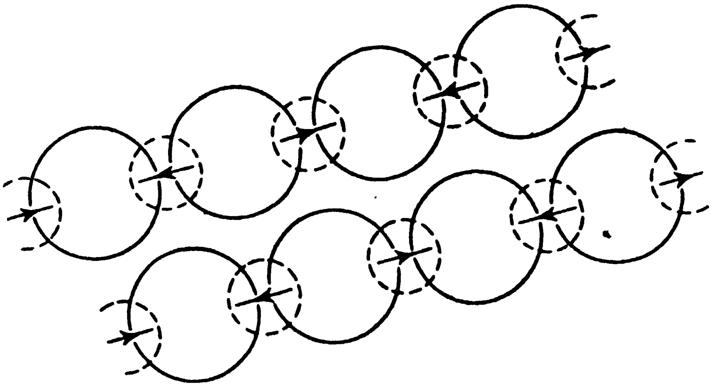


Fig. 1A.

$$\frac{1}{2} . H_c . I = \frac{1}{2} \times 6{\cdot}5 \times 10^6 \times 1700 = 5{\cdot}5 \times 10^9 \text{ ergs}$$

and the internal stress about 5500 atmospheres.

In nickel the intrinsic pressure is $1{\cdot}4 \times 10^9$ dynes per square centimetre or 1400 atmospheres; in cobalt $4{\cdot}4 \times 10^9$ dynes per square centimetre or 4400 atmospheres, in cast iron 4×10^9 dynes per square centimetre or 4000 atmospheres. These internal stresses are a measure of the forces binding the molecules together and should give an estimate of the ultimate tensile strength or tenacity of the medium. Moreover the tenacities of ferro-magnetic, paramagnetic and diamagnetic media should be roughly of the same order. That this is so is seen from the following values† :—

	Material.	Tenacity (dynes per square centimetre).
Ferro-magnetic	Iron, (wire)	$5{\cdot}0-6{\cdot}0 \times 10^9$
	Iron, wrought	$2{\cdot}9-4{\cdot}5 \times 10^9$
	Iron, cast	$1{\cdot}2-1{\cdot}9 \times 10^9$
	Nickel (wire)	$5{\cdot}3 \times 10^9$
	Mild steel (0·2 per cent. carbon) .	$4{\cdot}3-4{\cdot}9 \times 10^9$
	High carbon steel	$7{\cdot}0-7{\cdot}7 \times 10^9$
	Nickel steel (5 per cent.)	$6{\cdot}2 \times 10^9$

---

* This, as we have seen in Part III., p. 93, is compatible with the values of the latent heat of fusion of diamagnetic crystalline media. See also *infra*, pp. 253–4, p. 256.

† KAYE and LABY, 'Physical and Chemical Constants,' 1918, p. 28.

	Material.	Tenacity (dynes per square centimetre).
Diamagnetic. .	Lead	$0{\cdot}16 \times 10^{9}$
	Zinc	$1{\cdot}1 - 1{\cdot}5 \times 10^{9}$
	Glass (sometimes paramagnetic) .	$0{\cdot}3 - 0{\cdot}9 \times 10^{9}$
	Quartz	10×10^{9}
	Copper	$1{\cdot}2 - 2{\cdot}5 \times 10^{9}$
	Silver (wire)	$2{\cdot}9 \times 10^{9}$
	Gold (wire)	$2{\cdot}6 \times 10^{9}$
Paramagnetic .	Aluminium	$0{\cdot}6 - 1{\cdot}5 \times 10^{9}$
	Tin	$0{\cdot}16 - 0{\cdot}38 \times 10^{9}$
	Glass (sometimes diamagnetic) . .	$0{\cdot}3 - 0{\cdot}9 \times 10^{9}$
	Platinum (wire)	$3{\cdot}3 \times 10^{9}$

We may conclude that the intrinsic stress due to the local molecular field, calculated as above, gives a satisfactory interpretation of the ultimate tensile strength of crystalline media, as observed experimentally, irrespective of the nature of their magnetic property.

The energy associated with the medium, in virtue of this intrinsic pressure of crystallization, is consistent to the right order with the value of the latent heat of fusion of the material.*

The energies per cubic centimetre of various diamagnetic and paramagnetic media, calculated from the latent heats of fusion, are given below :—

1. ORGANIC Compounds Investigated in Parts I. and III.

Substance.	Magnetic nature.	Latent heat.	Density.	Energy (ergs per cubic centimetre).
Benzene	diamagnetic	30 (calories per gramme)	0·88	$1{\cdot}1 \times 10^{9}$
Xylene	,,	39	—	—
Chlorobenzene . . .	,,	30	1·12	$1{\cdot}4 \times 10^{9}$
Bromobenzene . . .	,,	20	1·49	$1{\cdot}3 \times 10^{9}$
Aniline	,,	21	1·02	$0{\cdot}9 \times 10^{9}$
Acetophenone . . .	,,	33	—	—
Benzophenone . . .	,,	23	1·10	$1{\cdot}0 \times 10^{9}$
Phenylhydrazine . .	,,	36	1·10	$1{\cdot}6 \times 10^{9}$
Pyridine	,,	22	0·985	$0{\cdot}9 \times 10^{9}$
Nitrobenzene . . .	,,	22	1·19	$1{\cdot}1 \times 10^{9}$
Naphthalene. . . .	,,	35	1·15	$1{\cdot}7 \times 10^{9}$
Naphthylamine. . .	,,	22	—	—
Acetic acid	,,	44	1·05	$1{\cdot}9 \times 10^{9}$
Glycerine.	,,	42	1·26	$2{\cdot}2 \times 10^{9}$
Carbon tetra-chloride .	,,	4	1·58	$0{\cdot}26 \times 10^{9}$

---

* See also Part III., p. 93.

2. ELEMENTS and Inorganic Compounds.

Substance.	Magnetic nature.	Latent heat.	Density.	Energy (ergs per cubic centimetre).
Bismuth	diamagnetic	13 (calories per gramme)	9·8	$5 \cdot 3 \times 10^9$
Cadmium	„	14	8·6	$5 \cdot 0 \times 10^9$
Lead	„	5	11·4	$2 \cdot 4 \times 10^9$
Silver	„	22	10·5	$9 \cdot 7 \times 10^9$
Zinc	„	28	7·1	$8 \cdot 3 \times 10^9$
Phosphorus	„	5	1·8	$0 \cdot 38 \times 10^9$
Mercury	„	3	13·6	$1 \cdot 7 \times 10^9$
Copper	„	43	8·9	$16 \cdot 0 \times 10^9$
Sulphur	„	9	2·0	$0 \cdot 8 \times 10^9$
Ice	„	80	0·9	$3 \cdot 0 \times 10^9$
Aluminium	paramagnetic	77	2·7	$8 \cdot 7 \times 10^9$
Tin	„	14	7·3	$4 \cdot 3 \times 10^9$
Palladium	„	36	11·4	$17 \cdot 0 \times 10^9$
Platinum	„	27	21·5	$24 \cdot 0 \times 10^9$
Potassium	„	16	0·86	$0 \cdot 56 \times 10^9$
Iron* (transformation $\beta - \gamma$ at A_3 point) .	--	1·4 (calories per gramme)	7·9	$0 \cdot 46 \times 10^9$

In the case of iron at the A_3 point, the transformation is from one cubic crystalline form to another, and we should expect the change of internal energy to be smaller than in the general case of actual crystallization from the liquid state.

We can obtain a measure of intrinsic pressures in crystalline media in another way, which depends on extrapolation of the relation connecting the temperature of the freezing point with applied pressure.

If

v_l = volume of 1 gramme of liquid,
v_c = volume of 1 gramme of crystal,
ϑ = temperature of fusion,
p = pressure, in atmospheres,
L = latent heat,

we know that

$$\frac{\partial \vartheta}{\partial p} = \frac{(v_l - v_c) . \vartheta}{L} = \frac{\delta V . \vartheta}{L} .$$

If the applied pressure be such that $v_l = v_c$ then $\delta V = 0$, and if we can determine the pressure π for which this condition exists, we have determined the intrinsic pressure due to the crystalline grouping, for if the latter were greater than, or less than, π there would be a change of volume on crystallization. If $v_l = v_c$ since, as the curve of fusion shows, both ϑ and L are finite, $\frac{\partial \vartheta}{\partial p}$ must be zero, *i.e.*, we have to determine the value of the applied pressure corresponding to the maximum of the closed area on the ϑ, p diagram. We cannot expect this method to give us anything but an approximate value of π because the extrapolation beyond laboratory pressures is considerable, but the results of the calculation are suggestive.

---

* A. E. OXLEY, 'Trans. Faraday Society,' vol. XI., Part 2, February, 1916.

The relations between the temperature of fusion (ϑ) and applied pressure (p) in atmospheres, for the substances* here referred to are taken from 'Kristallisieren und Schmelzen' by G. TAMMANN, Leipzig, 1903, p. 204, *et seq.*

For *water* TAMMANN found

$$\vartheta-22 = 0{\cdot}00438\,.\,(p-2200)-77\times10^{-8}\,.\,(p-2200)^2.$$

Differentiating and equating $\frac{\partial\vartheta}{\partial p}$ to zero,

$$\frac{\partial\vartheta}{\partial p} = 0{\cdot}00438-154\times10^{-8}\,.\,(p-2200) = 0,$$

or

$$\pi = \frac{0{\cdot}00438+154\times2200\times10^{-8}}{154\times10^{-8}}$$

$$\doteqdot 5000 \text{ atmospheres} \doteqdot 5\times10^9 \text{ dynes/square centimetre.}$$

Benzophenone—

$$\vartheta = 48{\cdot}11+0{\cdot}02757p-0{\cdot}00000136\,.\,p^2.$$

Differentiating and equating $\frac{\partial\vartheta}{\partial p}$ to zero we find the intrinsic pressure

$$\pi \doteqdot 10{,}000 \text{ atmospheres.}$$

Acetophenone—

$$\vartheta = 19{\cdot}2+0{\cdot}0235p-0{\cdot}00000152\,.\,p^2$$

$$\pi \doteqdot 7700 \text{ atmospheres.}$$

Aniline—

$$\vartheta = -6{\cdot}1+0{\cdot}0203p-0{\cdot}00000112\,.\,p^2$$

$$\pi \doteqdot 9000 \text{ atmospheres.}$$

Nitrobenzene—

$$\vartheta = 5{\cdot}67+0{\cdot}02344p-0{\cdot}00000116\,.\,p^2$$

$$\pi \doteqdot 10{,}000 \text{ atmospheres.}$$

Xylene—

$$\vartheta = 13{\cdot}2+0{\cdot}03438p-0{\cdot}00000171\,.\,p^2$$

$$\pi \doteqdot 10{,}000 \text{ atmospheres.}$$

Benzene—

$$\vartheta = 5{\cdot}43+0{\cdot}0283p-0{\cdot}00000198p^2$$

$$\pi \doteqdot 7100 \text{ atmospheres.}$$

---

* Most of these substances show a change of diamagnetic susceptibility on crystallization. See Part I., pp. 120–131; Part III., pp. 96–97.

Naphthalene—

$$\vartheta = 79{\cdot}95+0{\cdot}03657p-0{\cdot}00000180p^2$$
$$\pi \doteqdot 10{,}100 \text{ atmospheres.}$$

Carbon tetra-chloride—

$$\vartheta = -23{\cdot}0+0{\cdot}0350p-0{\cdot}00000147p^2$$
$$\pi \doteqdot 11{,}900 \text{ atmospheres.}$$

Ethylene di-bromide—

$$\vartheta = 9{\cdot}85+0{\cdot}0252p-0{\cdot}00000125p^2$$
$$\pi \doteqdot 10{,}000 \text{ atmospheres.}$$

Formic acid—

$$\vartheta = 7{\cdot}75+0{\cdot}01276p-0{\cdot}00000080p^2$$
$$\pi \doteqdot 7980 \text{ atmospheres.}$$

Potassium—

$$\vartheta = 59{\cdot}5+0{\cdot}0146\ -0{\cdot}0000007p^2$$
$$\pi \doteqdot 10{,}000 \text{ atmospheres.}$$

Phosphorus—

$$\vartheta = 43{\cdot}93+0{\cdot}0275p-0{\cdot}00000050p^2$$
$$\pi \doteqdot 27{,}500 \text{ atmospheres.}$$

Sulphur, rhombic-monoclinic[*]

$$\vartheta = 95{\cdot}4+0{\cdot}03725p+0{\cdot}00000213p^2$$
$$\pi \doteqdot -8700 \text{ atmospheres.}$$

Solid CO_2—

$$\vartheta = -56{\cdot}8+0{\cdot}01999p-0{\cdot}000000755p^2$$
$$\pi \doteqdot 13{,}300 \text{ atmospheres.}$$

These values, it is true, appear rather high when compared with those found from other considerations, but due importance must be attached to the difficulties of experimental work of this nature and to the fact that the experimental data have been extrapolated over a considerable pressure interval (several thousands of atmospheres).

Comparing these results with those given on pp. 252 to 254 it is considered that a mean value of the intrinsic stress in diamagnetic crystalline media, viz., 2×10^9 dynes per square centimetre is representative of the true order of magnitude of the forcive which binds the molecules in the space lattice of a crystalline medium. This implies that the energy per unit volume of the diamagnetic crystalline medium, in virtue of the crystalline grouping, is comparable with 2×10^9 ergs.

---

[*] This transition of sulphur from the rhombic to the monoclinic form is accompanied by thermal absorption. Since $\frac{\partial^2\vartheta}{\partial p^2}$ is positive the transition line for different pressures will be convex to the p-axis.

As the corresponding stresses and energies are of the same order in the ferro-magnetic metals, and, further, since it has been shown (see Part II., pp. 143, 145 and Part III., pp. 84–87) that the local molecular forcive in diamagnetic media is of the same order as that in the ferro-magnetic metals, we may conclude that the local intensities of magnetization in the two types of media are comparable.

Since

$$\tfrac{1}{2} . H_c . I \doteqdot 2 \times 10^9 \text{ ergs},$$

and

$$H_c \doteqdot 10^7 \text{ gauss (see Part III., p. 86)},$$

we find

$$I \doteqdot 400,$$

and

$$a'_c = \frac{H_c}{I} \doteqdot 2{\cdot}5 \times 10^4.$$

In a diamagnetic crystalline medium the local forcives are comparable with those in iron, and, since the latter medium shows hysteresis in a magnetic field, we may enquire whether a similar phenomenon will be shown by diamagnetic media. If the diamagnetic molecules are magnetically unsymmetrical, the application of an external magnetic field will tend to orientate them.* But this will be a differential effect on our conception of a diamagnetic molecule, and thus the tendency of the applied field to produce new molecular groupings will be small. We should therefore expect that hysteresis due to magnetization will be inappreciable in diamagnetic media. In iron, on account of the unbalanced magnetic nature of the molecules or atoms, new groupings are actually produced under fields of moderate intensity and the formation of these implies a loss of energy which is measured by the area of the hysteresis loop.

If, however, we take a diamagnetic copper wire and subject it to mechanical strain, the medium shows mechanical hysteresis. If sufficiently large stresses are employed, a permanent set is produced within the individual crystalline grains, new groupings of the molecules are formed, and a certain amount of energy is dissipated. All media, whether they are ferro-, para-, or diamagnetic, will show mechanical hysteresis. The difference from a magnetic point of view lies merely in the compensated nature of the diamagnetic molecule as compared with the uncompensated nature of the ferro-magnetic molecule, but the local forcives are comparable, so that under mechanical stress the mechanical hysteresis effects will be comparable.

(3) A COMPARISON OF THE ELASTICITIES OF SOME DIAMAGNETIC CRYSTALS WITH THOSE OF CRYSTALLINE PARAMAGNETIC AND FERRO-MAGNETIC MEDIA.

It is well known that the application of an external magnetic field alters the distribution of stress in a mass of iron crystals. On our theory we see how the molecules orientate themselves during crystallization under the influence of the local forcives which are characteristic of the molecular configuration. The influence of these local forcives will produce in the crystalline medium a distribution of internal stress which will in general be different across different planes, and in this way the planes of cleavage can be defined. In the direction where the stress is greatest, we should expect the elastic properties of the crystal to be abnormally high, comparable in fact with the elastic properties of steel.

---

* This orientation in a diamagnetic liquid gives rise to the induced magnetic double refraction. See Part III., p. 87.

In other directions we should expect the elastic properties to be less pronounced, and indeed the shearing of crystals, merely by the insertion of a knife blade and the application of small pressure parallel to a plane of cleavage, is evidence of this.

The following values of YOUNG's Modulus of Rigidity for various ferro-magnetic, paramagnetic, and diamagnetic media show that the power to resist distortion is of the same order whatever the magnetic nature of the crystalline medium.

	Substance.	Magnetic nature.	YOUNG's modulus (dynes per square centimetre).	Rigidity (dynes per square centimetre).
*	Iron (0·1 per cent. carbon)	ferro-magnetic	$2\cdot13 \times 10^{12}$	$8\cdot3 \times 10^{11}$ (calc.)
	Steel (1 per cent. carbon) .	„	$2\cdot09 \times 10^{12}$	$8\cdot1 \times 10^{11}$
	Nickel	„	$2\cdot02 \times 10^{12}$	$7\cdot7 \times 10^{11}$ (calc.)
	Aluminium	paramagnetic	$7\cdot05 \times 10^{11}$	$2\cdot67 \times 10^{11}$
	Tin	„	$5\cdot43 \times 10^{11}$	$2\cdot0 \times 10^{11}$ (calc.)
	Glass	(sometimes diamagnetic)	$6\cdot5 - 7\cdot8 \times 10^{11}$	$2\cdot6 \times 10^{11}$
	Platinum	paramagnetic	$1\cdot68 \times 10^{12}$	$6\cdot1 \times 10^{11}$
	Lead	diamagnetic	$1\cdot62 \times 10^{11}$	$0\cdot56 \times 10^{11}$ (calc.)
	Zinc	„	$1\cdot25 \times 10^{12}$	5×10^{11}
	Copper	„	$1\cdot23 \times 10^{12}$	$4\cdot55 \times 10^{11}$
	Silver	„	$7\cdot9 \times 10^{11}$	$2\cdot87 \times 10^{11}$
	Gold	„	$8\cdot0 \times 10^{11}$	$2\cdot77 \times 10^{11}$
	Bismuth	„	$3\cdot19 \times 10^{11}$	$1\cdot2 \times 10^{11}$ (calc.)
	Quartz (fibre)	„	$5\cdot18 \times 10^{11}$	$3\cdot0 \times 10^{11}$
			Principal YOUNG's moduli (dynes per square centimetre).	Principal rigidity (dynes per square centimetre).
†	Quartz (crystalline) . . .	diamagnetic	$1\cdot00 \times 10^{12}$	$5\cdot7 \times 10^{11}$
	Beryl	paramagnetic	$2\cdot06 \times 10^{12}$ $2\cdot25 \times 10^{12}$	$6\cdot54 \times 10^{11}$ $9\cdot6 \times 10^{11}$
	Topaz	diamagnetic	$2\cdot25 \times 10^{12}$ $2\cdot83 \times 10^{12}$ $2\cdot60 \times 10^{12}$	— — —
	Rock salt	„	$0\cdot41 \times 10^{12}$	$1\cdot27 \times 10^{11}$
	Potassium chloride . . .	„	$0\cdot36 \times 10^{12}$	$0\cdot64 \times 10^{11}$
	Fluor spar	„	$1\cdot44 \times 10^{12}$	$3\cdot4 \times 10^{11}$
	Pyrites	—	$3\cdot46 \times 10^{12}$	$10\cdot5 \times 10^{11}$

---

* KAYE and LABY, 'Physical and Chemical Constants,' p. 27, 1918.

† A. E. H. LOVE, "The Mathematical Theory of Elasticity," 'Camb. Univ. Press,' p. 160. In the cases of beryl and topaz the different values correspond to bars whose lengths are in the directions of the different axes of symmetry.

As **Love** points out, the values of these elastic constants for beryl and topaz are remarkable in that they are greater than the corresponding constants in ordinary steel. The values of the elastic coefficients for most of the other substances in the above table are comparable with the constants for steel, and it is considered that these results give very strong evidence in favour of the large intermolecular forcive operative in diamagnetic crystalline media and confirm the suggestion made in Part II., p. 143, that this local forcive is comparable with that in ferro-magnetic media.

If a crystalline medium be heated, then as long as the crystalline state prevails, rotational vibrations of large amplitude are prevented, so that the specific heat of the crystalline medium is lower than that of the supercooled liquid.* In the latter case, the liquid at low temperatures passes into a rigid gel, and when this is heated, the molecules acquire rotational vibrations gradually until finally the ordinary liquid state is reached, possessing no appreciable rigidity. It is important to note that the molecules are vibrating under a local forcive to which we are ascribing the elastic properties of the medium, and therefore the theory is consistent with the theory of specific heat developed by **Debye**, in which the forces which control the thermal vibrations of the molecules are identical with those which determine the elastic constants of the medium. **Madelung** and **Sutherland** have similarly suggested that the elastic forces resisting mechanical strain are just those forces which determine the infra-red optical vibrations of the atoms in the solid substance. It has been found possible to calculate the infra-red frequencies from a knowledge of the mechanical properties. In the present researches it has been shown that we can calculate both the optical frequencies† and the mechanical stresses from the local molecular forcive. *Within* the core of the atom the local controlling force may be more intense, and although such an intense forcive would not be directly operative in determining the state of crystallization, yet it might be responsible for determining frequencies on the ultra-violet side comparable with X-ray frequencies. (See *infra*, pp. 273 and 278.)

(4) THE CHANGE OF DENSITY ON CRYSTALLIZATION INTERPRETED AS A MAGNETO-STRICTION EFFECT OF THE MOLECULAR FIELD.

If we subject a liquid to a magnetic field, a change of volume occurs to such an extent that the internal pressure is reduced by an amount equal to the potential energy per unit volume of the magnetic field. This change of internal pressure (see Part III., p. 91) is

$$\tfrac{1}{2} \cdot k_l \cdot \mathrm{H}^2 + \tfrac{1}{2} \cdot \lambda \cdot k_l^2 \cdot \mathrm{H}^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (2)$$

where

k_l is the susceptibility of the liquid per unit volume,

λ, a constant equal to 1/3,

and

H, the applied field intensity.

---

* Part III., p. 94.

† A. E. **Oxley**, 'Roy. Soc. Proc.,' A, vol. 95, p. 58, 1918, and Part III., p. 84.

As k_l for diamagnetic media is of the order -7×10^{-7}, the second term in (2) is insignificant in comparison with $\frac{1}{2}\,.\,k_l\,.\,H^2$.

If c be the compressibility of the liquid, the change of volume in cubic centimetres per cubic centimetre will be given by

$$\delta v = \tfrac{1}{2}\,.\,c\,.\,k_l\,.\,H^2 \quad \text{. . . . .} \quad (3)$$

This relation has been experimentally verified by QUINCKE.* The compressibility c is of the order 10^{-10}, particular values for different substances being :—

Substance.	$c\times 10^{10}$.
Benzene	0·8
Chlorobenzene	0·7
Toluene	0·8
Xylene	0·7
Water	0·5
Carbon tetrachloride	0·9
Acetic acid	0·4
Carbon bisulphide	0·9

Substance.	$c\times 10^{12}$.
Mercury	3·7
Potassium	31·5
Sodium	15·4
Lead	2·2
Tin	1·7
Bismuth	2·8
Iron	0·4

Since the largest magnetic field at our disposal is 50,000 gauss, the largest value of δv is

$$-\tfrac{1}{2}\times 10^{-10}\times 7\times 10^{-7}\times 2{\cdot}5\times 10^{9} = -8\times 10^{-8} \doteqdot -10^{-7}\ \text{c.c./c.c.}$$

Now we have shown (Part III., p. 90) that the potential energy term corresponding to (2) for a crystalline medium is

$$\tfrac{1}{2}\,.\,k_l\,.\,H^2+\tfrac{1}{2}\,.\,a'_c\,.\,I^2$$

per unit volume, where a'_c is the constant of the local molecular field and I is the aggregate of the local intensity of magnetization per unit volume. The term $\frac{1}{2}\,.\,a'_c\,.\,I^2$ is associated with each cubic centimetre of the crystalline structure whether

---

* See G. T. WALKER, "Aberration and the Electromagnetic Field," 'Camb. Univ. Press,' pp. 72–83.

H be zero or not. It is really due to the spontaneous *local* intensity of magnetization per unit volume and corresponds to the similar energy term, $\frac{1}{2}NI^2$ (WEISS), due to the spontaneous magnetization in iron. Using the values which have already been assigned to a'_c and I,[*] viz.,

$$\left.\begin{array}{l} a'_c \doteqdot 2{\cdot}5\times 10^4\,\text{[†]} \\ I \doteqdot 400 \end{array}\right\}, \qquad (4)$$

we see at once that this term is large in comparison with $\frac{1}{2}\,.\,k_l\,.\,H^2$ and therefore the change of volume to which the potential energy term $\frac{1}{2}\,.\,a'_c\,.\,I^2$ gives rise will be large in comparison with that which we can produce artificially in a liquid by applying the largest field available in the laboratory.

When a diamagnetic substance crystallizes, the alteration of internal pressure will be $\frac{1}{2}a'_c\,.\,I^2$ and therefore the accompanying change of volume in cubic centimetres per cubic centimetre will be

$$\delta V = \tfrac{1}{2}\,.\,c\,.\,a'_c\,.\,I^2,$$

and substituting from (4) we find

$$\delta V = \tfrac{1}{2}\,.\,0{\cdot}8\times 10^{-10}\times 16\times 10^4\times 2{\cdot}5\times 10^4 = 0{\cdot}16 \text{ c.c./c.c.}$$

to the appropriate order.

Some values of δV in cubic centimetres per cubic centimetre are :—

Substance.	δV.	
Benzene	0·10	[‡]
Naphthalene	0·14	
Benzophenone	0·19	
Di-phenylamine	0·10	
Formic acid	0·10	
Sodium	0·03	
Potassium	0·03	
Mercury	0·036	
Lead	0·03	[§]
Tin	0·03	
Bismuth	0·03	
Iron (at A_3 point)	0·003	

The values calculated agree as well as could be expected with the experimental determinations, since we know the orders of magnitude only of a'_c and I, for these are unknown functions of the molecular structure and space lattice of each substance. It will be noticed that those substances showing a small value of δV have low compressibilities.

---

[*] See p. 257 *supra*.

[†] A physical explanation of this large value of a'_c and of the corresponding constant N in ferromagnetism is given on p. 267 *infra*.

[‡] G. TAMMANN, 'Kristallisieren u. Schmelzen,' Leipzig, 1903, pp. 204 *et seq.*

[§] DESCH, 'Metallography,' p. 242.

The case of iron at the A_3 transformation is particularly interesting. From the curves given by CHARPY and GRENET,* dealing with the expansion of iron and iron-carbon alloys between 200° C. and 1000° C., we can show that the extent of linear contraction which occurs suddenly at the A_3 point (900° C.) is of the order 0·003 centimetre per centimetre. The change of volume will be of this order of magnitude, which is small in comparison with the change of volume accompanying the crystallization of many organic compounds, but is very large compared with the magneto-striction effect which can be induced in either a ferro-magnetic or diamagnetic substance with a field of 50,000 gauss.

Taking into account the small compressibilty of iron,† which is only $0{\cdot}4 \times 10^{-12}$, or about 1/200 that of the liquids above referred to, this change of volume may be interpreted as due to a change of internal energy represented by $\frac{1}{2} . N . I^2$ where N is the constant of the ferro-magnetic field, of the order $0{\cdot}38 \times 10^4$,‡ I the saturation intensity of magnetization, of the order 1760. For we have

$$\begin{aligned}\delta V &= \tfrac{1}{2} . c . N . I^2 \\ &= \tfrac{1}{2} \times 0{\cdot}4 \times 10^{-12} \times 0{\cdot}38 \times 10^4 \times 1{\cdot}76^2 \times 10^6 \\ &= 0{\cdot}002\,(4) \text{ c.c./c.c.}\end{aligned}$$

which is of the order of magnitude found experimentally. The molecular field exists in an unmagnetized piece of iron and is accompanied by the large spontaneous magnetization of that element throughout an individual grain, but as these grains have all types of orientation, the large molecular field and the accompanying spontaneous magnetization are hidden in a piece of iron large enough to contain many grains. The molecular field will nevertheless produce the magneto-striction effect referred to above.§ Let us suppose that such a piece of iron is subjected to an external magnetic field. The molecules of all the grains will tend to come into alignment with the applied field and there will be a new distribution of stress. In an unmagnetized piece of iron, taken as a whole, the stress may be regarded as equal in all directions, but when an external field is applied, this is no longer the case and the iron shows a new magneto-striction effect consisting of an expansion in one direction and a contraction in the other. The extent of the redistribution of stress should be determined by a term of the form $\frac{1}{2} . a'_c . i^2$ where a'_c is the constant of the molecular field and ι is the resultant intensity of magnetization induced in the direction of the external field. The compressibilities will now be different along and perpendicular to this direction, and the change of volume should be proportional to ι^2 for the particular value of the applied field. This result agrees with experiment.*

---

* *Loc. cit.*

† RICHARDS, 'Journ. Chem. Soc.,' vol. 99, p. 1201, 1911.

‡ WEISS and BECK, 'Journ. de Phys.,' sér. iv., vol. 7, p. 249, 1908.

§ See *infra*, p. 265. It is assumed here that the molecular field disappears just above the A_3 point, at least in so far as it is effective in causing spontaneous magnetization. This is in accordance with the small paramagnetic susceptibility of iron above A_3.

We can see in a general way how the sign of δV on crystallization may sometimes be positive and sometimes negative. Usually the molecular packing in the crystalline state will be closer than that in the liquid state, but it may happen that the configuration of the molecule is such that, when the parts which have the strongest magnetic attraction for each other are in the position of minimum potential energy, the packing is more open than in the liquid state. The appearance of the internal forcive on crystallization will in this case be accompanied by expansion. In the former case the appearance of the internal forcive will be accompanied by contraction.

The region of stability of the crystalline state is represented on the pressure temperature diagram by a closed area (fig. 2).

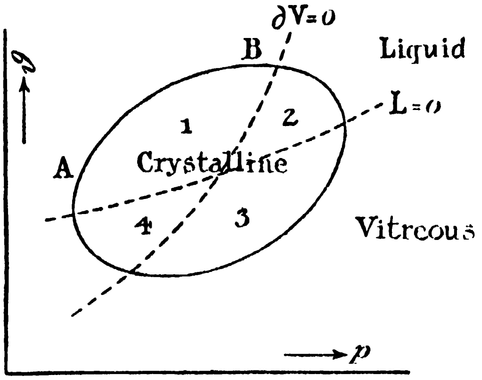


Fig. 2.

This area will in general be divided into four quadrants by the loci of the lines $\delta V = 0$ and $L = 0$, where δV is the change of volume on crystallization and L is the latent heat.

In the four quadrants the following conditions hold :—

Quadrant.	L.	∂V.	$\frac{\partial \vartheta}{\partial p}$.	$\frac{\partial^2 \vartheta}{\partial p^2}$.
1	+	+	+	−
2	+	−	−	−
3	−	−	+	+
4	−	+	−	+

If the melting-point is at some point along the arc AB of the first quadrant, increase of pressure raises the m.p. and δV and L are positive. This corresponds vith the case for benzophenone. In the case of water the diagram is as in fig. 3 and or $p = 1$ atmosphere, δV is negative and L is positive. The m.p. is in this case ocated on the arc A′B′ in the second quadrant. A similar case is that of pure iron as he temperature is raised through the critical point A_3 (see fig. 4). At A_3, δV is negative and L is positive, while $\frac{\partial \vartheta}{\partial p}$ is negative.

---

* NAGAOKA and HONDA, 'Phil. Mag.,' vol. 46, p. 268, 1898.

Now

$$\frac{\partial \vartheta}{\partial p} = \frac{(v_\gamma - v_\beta) . \vartheta}{L} = \frac{\delta V . \vartheta}{L}.$$

Therefore $v_\beta > v_\gamma$ which gives a shrinkage in iron on heating through the A_3 point at bout 900° C.

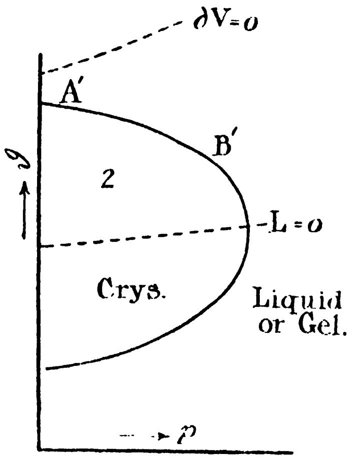


Fig. 3.

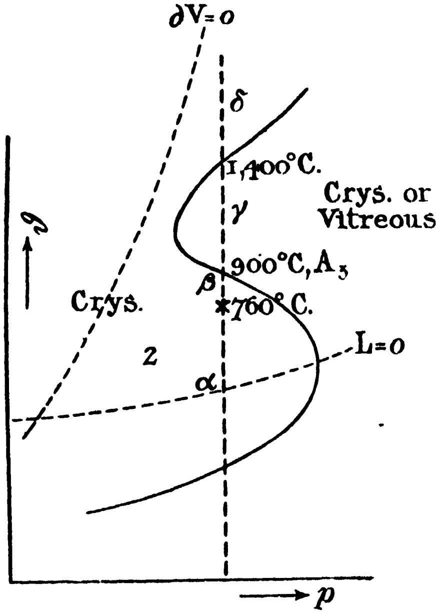


Fig. 4.

The application of the relation

$$\frac{\partial \vartheta}{\partial p} = \frac{(v' - v'') . \vartheta}{L} = \frac{\delta V . \vartheta}{L}$$

s interesting in connexion with some abnormalities of heats of recalescence in ferronagnetic media. The change of volume δV, both expansion and contraction; the absorption or evolution of heat of the amount L, depend solely upon the shape of the region of crystalline stability and its position relatively to the ϑ, p axes.

Thus nickel steel with no carbon and pure cobalt show no recalescence at the nagnetic change points. In these cases L is very small and since δV and ϑ are finite, $\frac{\partial \vartheta}{\partial p}$ will be large (fig. 5). In other words, under the pressure of one atmosphere, the path of the crystallization curve cuts the line AB at a steep angle in the neighbourhood of the intersection of the fusion curve with the neutral line $L = 0$.

The apparent discontinuities of the susceptibility temperature curve are suggestive in this connexion (fig. 6). The upper branch of the curve AB for the so-called β-iron is practically continuous with the branch CD for the δ-range above 1400° C. Between B and C there is a break, the branch BE representing β-iron over a somewhat narrow range just below the critical temperature ϑ_c and the branch EF representing paramagnetic γ-iron. The locus BEFC corresponds to a crystalline modification of iron which is more stable over this temperature interval than the crystalline grouping or groupings over the ranges AB and CD.

If this view is correct, the molecular field is operative in iron over a temperature interval from 1400° C. upwards as well as below the critical temperature. The existence of this forcive above 1400° C. implies a crystalline symmetry involving an appreciable mutual action between the molecules consistent with the enhanced susceptibility found by CURIE* and by WEISS and FOEX† over this range. In the intermediate range, between 900° C. and 1400° C., iron shows a paramagnetic quality only. Perhaps we may regard the molecular state in this range as more allied to a gel, consisting of very small interlocked grains, each with relatively few molecules,‡ rather than to a coarse grain crystalline arrangement of the molecules. The orientations of the molecular axes as we pass from one small grain to another will be different, so that each grain is, as it were, surrounded by a surface of vitreous material. As stated on p. 251 the molecular field would then be non-effective in so far as the production of

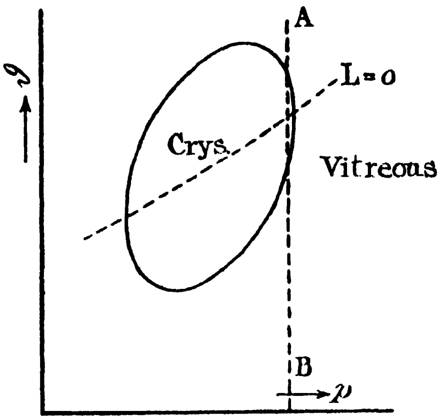


Fig. 5.

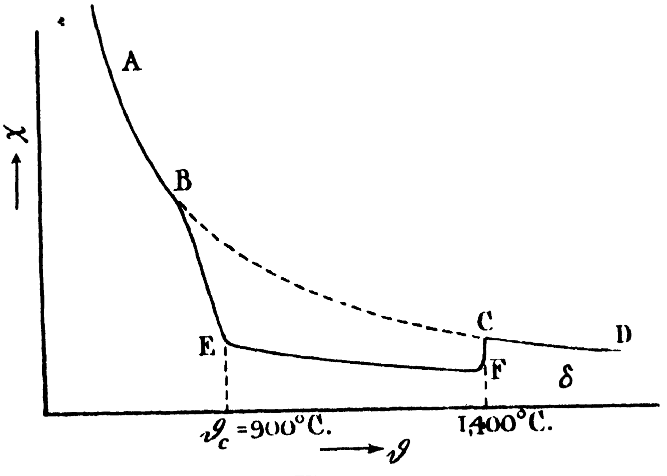


Fig. 6.

---

* 'Annales de Chimie et de Physique,' sér. iv., vol. v., p. 289, 1895.

† 'Archives des Sciences, Genève,' sér. 4, vol. xxxi., p. 88, 1911.

‡ This smallness of grain structure above A_3 in wrought iron or mild steel is consistent with the experiments of J. E. STEAD, 'Iron and Steel Institute,' 1898, No. 1, p. 145. The reverse effect, that of a coarse structure on cooling from above A_3, observed by STEAD and CARPENTER ('Iron and Steel Institute,' No. 11, p. 119, 1913) in the case of thin strips of electrolytic iron, may possibly be attributed to surface forces. See also a paper "On the Part Played by the Amorphous Phase in the Hardening of Steels," by J. C. W. HUMFREY, 'Trans. Faraday Soc.,' May, 1915.

spontaneous magnetization is concerned, but the tenacity might even surpass that of the chaotic β-state since the transition $\beta-\gamma$ is accompanied by shrinkage. Such an increase of tenacity was actually found by **ROSENHAIN** and **HUMFREY**.*

The effects of prolonged heating on the plasticity of mild steel are interesting in this connexion. **EWING** and **ROSENHAIN**† have shown that the plasticity of a material is caused by slips occurring on cleavage or "gliding" surfaces *within each* of the crystalline grains, although the elementary portions of the crystals retain their primitive form and the crystalline structure of the metal as a whole is preserved. In the case of mild steel, exposure to a temperature of 1200° C. or higher temperatures for several hours may cause the material to lose much of its plasticity, while some specimens of soft iron after prolonged exposure at 700° C. to 800° C. (less than the critical temperature) have been made brittle. These results are consistent with a similarity of molecular configuration for temperatures above 1200° C. and below the critical temperature (about 850° C.). On the other hand an exposure at 900° C. or 1000° C. (*i.e.*, in the region of the paramagnetic state), followed by a slow or fairly rapid cooling, induces considerable plasticity in the material, and this treatment may even be used to remove brittleness originating from heating to the higher or lower ranges of temperature mentioned above. As the plasticity is produced by slipping on cleavage surfaces *within* the crystalline grains, this smallness of the grain structure may, under stress, determine a molecular rotation. **EWING** and **ROSENHAIN**‡ have shown that in some metals, in addition to slips or motions of pure translation, there results a molecular rotation from strain which gives rise to twin-crystals. It is interesting to note that the formation of twin-crystals is common in iron through the γ-range, but has not been observed in the β and α ranges.

As the temperature falls below 1400° C. there is some modification of the crystalline cubic arrangement, resulting in a closer packing of the molecules, and accompanied by an interlocking of the fine grains. Thus iron in the γ-range (1400° C. to 900° C.) will be paramagnetic. At a lower temperature than 900° C., this state is unstable and another modification of the crystalline grouping occurs, accompanied by expansion, thermal evolution, and the appearance of spontaneous magnetization. This latter effect seems inconsistent with a more open packing of the molecules, but an analogy is found in the case of water, where the molecular influence in the liquid just above 0° C. is small compared with that in ice just below freezing point, although the packing of the molecules in the liquid state is closer than that in the crystalline state (see *supra* p. 264). As the temperature is lowered the transformation progresses rapidly until a point B is reached, after which the increase of magnetic quality is somewhat less rapid. On continued cooling, the iron passes into the α-range where the magnetic property is capable of attaining a saturation value.

---

* 'Roy. Soc. Proc.,' A, vol. 83, 1909; 'Iron and Steel Institute,' No. 1, 1913.

† 'Phil. Trans. Roy. Soc.,' A, vol. 193, p. 279.

‡ See **EWING**, 'The Strength of Materials,' p. 47, 1906.

(5) FURTHER DISCUSSION OF THE NATURE OF THE LOCAL MOLECULAR FIELD IN FERRO-MAGNETIC AND DIAMAGNETIC MEDIA.

This subject has already been discussed in Part III., pp. 89 and 100, but no interpretation was then given to the magnitudes of the constants N and a'_c of the local ferro-magnetic and diamagnetic fields. The values of N given by WEISS and BECK* are $0·38 \times 10^4$ for iron, and $1·27 \times 10^4$ for nickel. For diamagnetic crystalline media a'_c is of the order 2×10^4.

EWING and LOW† have shown that in very strong magnetic fields the relation between induction (B) and applied field (H) may be represented by the equation

$$B = H + \text{a constant}. \qquad (1)$$

This constant has the value $4\pi I$ (where I is the saturation intensity) and is equal to 21,360 in wrought iron, 6470 in nickel and 16,300 in cobalt.

In the case of wrought iron

$$B = H + 21360. \qquad (2)$$

Suppose we could apply a field equal to the molecular field, $6·53 \times 10^6$ gauss for iron. The limiting value of the permeability μ_L for this field will be from (2)

$$\mu_L = 1 + \frac{21360}{6·53 \times 10^6},$$

and the limiting susceptibility per unit volume

$$\chi_L = \frac{\mu_L - 1}{4\pi} = \frac{21360}{4\pi \times 6·53 \times 10^6} = 2·60 \times 10^{-4},$$

and

$$\frac{1}{\chi_L} = 0·38 \times 10^4.$$

This is equal to the value of N, the coefficient of the molecular field, as we should expect.

Similar calculations may be made for nickel and cobalt, the limiting susceptibilities being respectively

$$\chi_L = 0·81 \times 10^{-4} \text{ for nickel,}$$

$$\chi_L = 2·0 \times 10^{-4} \text{ for cobalt.}$$

Now we may ask the question, why is it that, in spite of the fact that all the molecules are ordered into a definite space lattice under the influence of the respective molecular fields, the materials still show a finite susceptibility to magnetization? The explanation, I think, is to be found in the finite though small angular oscillations which constitute a portion of the thermal energy of the molecules. The molecules are fixed relative to one another and form a definite space lattice but they are oscillating with small amplitude under the molecular field. This allows them to retain a finite susceptibility. Suppose we could double the molecular field, the limiting susceptibility would become one-half its former value and the saturation intensity of magnetization would be slightly increased. If we could increase the molecular field indefinitely, the susceptibility would get indefinitely small, the product of the two however tending to a finite limit equal to the *true* saturation intensity of magnetization. The amplitude of the molecular oscillations would, under the influence of this indefinitely large forcive, be indefinitely small. This state might be attained in a practical manner by cooling the substance, say in liquid hydrogen, when the limiting susceptibility would become vanishingly small.

---

* 'Journ. de Phys.,' sér. iv., vol. 7, p. 249, 1908.

† 'Phil. Trans. Roy. Soc.,' A, p. 242, 1889.

As N is the reciprocal of the limiting susceptibility the constant of the molecular field will become indefinitely large. WEISS, however, supposes N to be constant.* The tendency of χ_L to approach a small limiting value as the temperature is lowered is confirmed experimentally for ferro-magnetic substances† and is particularly noticeable in the case of weak magnetic fields. The reduction of the amplitude of vibration of the molecules as the absolute zero is approached merely implies a higher frequency of angular oscillation under the increasing molecular field and does not necessarily imply that the rotational energy becomes vanishingly small. In this case it should be noted that the saturation intensity of magnetization we are considering is smaller than that which would be given by the simple summation of all the magnetic moments of the molecules in unit volume. In other words, the difficulty of producing this latter saturation by an external field becomes increasingly difficult on account of the larger molecular forcive at low temperatures, in agreement with the vanishingly small susceptibility referred to above. At higher temperatures the susceptibility to an external field is far greater ; the molecules are, as it were, helped over their difficulties with respect to the molecular field, when the external field is applied, by the increased energy of the rotational oscillations, and having passed this critical point they are held in new combinations. Beyond the critical point the molecular state is chaotic, the molecules being interlocked (*cf.* p. 265), and the external field has sufficient control to produce a paramagnetic effect only.

---

* Following WEISS, we have taken the molecular field proportional to I. WEISS writes the molecular field NI and assumes N to be constant. This applies with sufficient accuracy in a temperature region just below the critical temperature, but cannot be true over the whole region down to absolute zero, because, as the molecular translational vibrations die down, the molecules approach one another more closely and the molecular field must necessarily increase considerably although I remains practically constant. This increase is accounted for by the increase of the coefficient N, which is the reciprocal of the limiting susceptibility.

† EWING, 'Magnetic Induction in Iron and other Metals,' p. 172 *et seq.*, where curves are given for iron, hard steel, nickel and various nickel steels. See also p. 269 *infra* and EWING, *loc. cit.*, p. 354.

The fall off in the value of χ as the temperature is reduced, may, in part, be explained by the increased value of the molecular field, due to the nearer approach of the molecules. As the local molecular field becomes very big, the induced diamagnetic effect in each molecule will become big, in the same proportion, and this will tend to reduce the value of χ and make this quantity tend to a limiting value. If the molecular field is of the order 10^7 gauss at ordinary temperatures, we have seen that the ratio of the induced diamagnetic moment ΔM to the magnetic moment of the electron orbit M, is of the order 1/100. If the molecular field at low temperatures approaches 10^9 gauss, owing to closer proximity of the magnetic elements in neighbouring molecules, the diamagnetic effect would be comparable with the ferro-magnetic effect.* On our view this does not imply that diamagnetic substances should acquire a large diamagnetic susceptibility at very low temperatures. For taking the molecules in pairs, locally they are paramagnetic and the action of the local molecular field is to reduce this paramagnetic effect so that the local magnetic moment becomes smaller and the susceptibility to an external field tends to zero as in iron.

According to EWING† experiments carried out to test this effect have neither proved nor disproved this theory, probably because the external fields were not sufficiently strong. But during crystallization we are applying unconsciously to each molecular current a magnetic field 500 or 1000 times stronger than the largest field we can apply externally, and probably even greater local intensities are attained at low temperatures, since the interacting magnetic elements in adjacent molecules may be almost touching one another. The mutual induction and temperature effects combine to cause χ to approach the limit zero at the absolute zero or in very powerful external fields.

In diamagnetic media we have seen that the constant of the local molecular field a'_c (which corresponds to N in WEISS's ferro-magnetic field) is of the order $2{\cdot}5 \times 10^4$ and the reciprocal of this, viz., $+4 \times 10^{-5}$, is the order of magnitude of the *local positive limiting susceptibility* of a diamagnetic crystalline medium. (At ordinary temperatures the diamagnetic susceptibility per unit volume is of the order -10^{-6}). The parts of molecules adjacent to one another in a diamagnetic crystalline medium attract in a similar manner to the adjacent parts of molecules of a ferro-magnetic or paramagnetic medium.

In a ferro-magnetic medium, as the temperature is raised, the susceptibility increases up to a certain point just below the critical temperature and then falls off rapidly. The temperature controls the susceptibility in two ways; first, by helping the molecules to overcome the difficulties of orientation, produced by the neighbouring molecules, to a point just below the critical temperature; second, by overdoing this effect and by giving the molecules too much rotational energy, at the critical temperature and above, so that the susceptibility to magnetization falls very rapidly.

---

* In nickel the molecular field is $6 \cdot 3 \times 10^6$ gauss. An applied field of 10^6 gauss would make a substance as diamagnetic as bismuth have a saturation value equal to that of nickel.

† 'Magnetic Induction in Iron and other Metals,' p. 353.

In a diamagnetic crystalline medium, as the temperature is raised, the local positive susceptibility will obey a similar law, the temperature of fusion now corresponding to the critical temperature in the ferro-magnetic case. Although locally the relation between susceptibility and temperature is the same in the two cases, the effect passes unnoticed in the diamagnetic case because the molecule has a total zero magnetic moment. Nevertheless, the effect of temperature acts in its two antagonistic ways in diamagnetic as well as in ferro-magnetic media. When the temperature is above the melting point, the rotational energy of the molecules annuls the local forcive (liquid state), when it is very low the molecules become interlocked and cannot readjust themselves in a space lattice (gel state). There is an intermediate region of temperature where opportunity is offered for the tendency of self-orientation under the mutual local forcives to display itself, and over this range crystallization may take place. This intermediate temperature range defines the closed region of stability of the crystalline form on the pressure temperature diagram of equilibrium of the crystalline and amorphous states (see p. 263).

A discussion as to how far we may regard it as proved that the local molecular field in crystalline media is of magnetic nature was given in para. 8, of Part III. The conclusion reached was that the molecular field is certainly in part magnetic. It is possible to bring forward further evidence of the truth of this deduction. In some noteworthy researches* published by TYNDALL, as long ago as 1870, it was shown that magnetic properties of crystalline media bear a close relation to molecular aggregation. About 100 different crystals were examined and from the deportment of these, when subjected to a magnetic field, TYNDALL found that "if the arrangement of the component particles of any body be such as to present different degrees of proximity in different directions, then the line of closest proximity, other circumstances being equal, will be that chosen by the respective forces for the exhibition of their greatest energy. If the mass be magnetic this line will stand axial, if diamagnetic, equatorial."†

The exactness of the dependence of magnetic deportment on the position of cleavage planes is remarkably shown in these experiments. Whatever the crystal examined, it was found that the magnetic deportment disclosed accurate information of the planes of cleavage. TYNDALL describes the results of his important experiments in such elegant language that it may be permissible to quote some of them at length. Thus he continues :—" From this point of view, the deportment of the two classes of crystals represented by Iceland spar and carbonate of iron, presents no difficulty. This crystalline form is the same, and as to the arrangement of the particles, what is true of one will be true of the other. Supposing then, the line of closest proximity to coincide with the optic axis; this line, according to the principle expressed, will stand axial or equatorial, according as the mass is magnetic or diamagnetic, which is precisely what the experiments with these crystals exhibit. Analogy as we have seen justifies the assumption here made. It will, however, be of interest to enquire, whether any discoverable circumstance connected with crystalline structure exists upon which the difference of proximity depends and knowing which, we can pronounce with tolerable certainty, as to the position which the crystal will take up in the magnetic field.

---

* 'On Diamagnetism and Magnecrystallic Action,' 1870.

† *Loc. cit.*, p. 23.

"The following experiments will perhaps suggest a reply.

"If a prism of sulphate of magnesia be suspended between the poles with its axis horizontal, on exciting the magnet the axis will take up the equatorial position. This is not entirely due to the form of the crystal; for even when its axial dimension is shortest, the axis will assert the equatorial position, thus behaving like a magnetic body, setting its longest dimension from pole to pole.

"Suspended from its end with its axis vertical, the prism will take up a determinate oblique position. When the crystal has come to rest, let that line through the mass which stands exactly equatorial be carefully marked. Lay a knife-edge along this line, and press it in the direction of the axis. The crystal will split before the pressure, disclosing shining surfaces of cleavage. This is the only cleavage the crystal possesses and it stands equatorial. Sulphate of zinc is of the same form as sulphate of magnesia, and its cleavage is discoverable by a process exactly similar to that just described. Both crystals set their planes of cleavage equatorial. Both are diamagnetic.

"Let us now examine a magnetic crystal of similar form. Sulphate of nickel is, perhaps, as good an example as we can choose. Suspended in the magnetic field with ts axis horizontal, on exciting the magnet the axis will set itself from pole to pole, and this position will be persisted in, even when the axial dimension is shortest. Suspended from its end, the crystalline prism will take up an oblique position with considerable energy. When the crystal thus suspended has come to rest, mark the ine along its end which stands *axial.* Let a knife edge be laid along this line and pressed in a direction parallel to the axis of the prism. The crystal will yield before the edge and discover a perfectly clean plane of cleavage.

"These facts are suggestive. The crystals here experimented with are of the same outward form; each has but one cleavage, and the position of this cleavage with regard to the form of the crystal, is the same in all. The magnetic force, however, at once discovers a difference of action. *The cleavages of the diamagnetic specimens stand equatorial; of the magnetic, axial.*

"A cube cut from a prism of scapolite, the axis of the prism being perpendicular to wo of the parallel faces of the cube, suspended in the magnetic field, sets itself with he axis of the prism from pole to pole.

"A cube of beryl of the same dimensions with the axis of the prism from which it is aken also perpendicular to two of the faces, suspended as in the former case, sets itself with the axis equatorial. Both these crystals are magnetic.

"The former experiments showed a dissimilarity of action between magnetic and diamagnetic crystals. In the present instances, both are magnetic, but still there is a difference; the axis of the one prism stands axial, the axis of the other equatorial. With regard to the explanation of this, the following fact is significant. Scapolite cleaves *parallel* to its axis, while beryl cleaves *perpendicular* to its axis; the cleavages in both cases, therefore, stand axial, thus agreeing with sulphate of nickel. The cleavages hence appear to take up a determinate position regardless of outward form, and they seem to exercise a ruling power over the deportment of the crystal.

"A cube of saltpetre, suspended with the crystallographic axis horizontal, sets itself between the poles with this axis equatorial.

"A cube of topaz, suspended with the crystallographic axis horizontal, sets itself with this axis from pole to pole.

"We have here a kind of complementary case to the former. Both these crystals are diamagnetic. Saltpetre cleaves parallel to its axis; topaz perpendicular to its axis. The planes of cleavage, therefore, stand in both cases equatorial, thus agreeing with sulphate of zinc and sulphate of magnesia.

"Where do these facts point? A moment's speculation will perhaps be allowed us here. May we not suppose these crystals to be composed of layers indefinitely thin, laid side by side, within the range of cohesion, which holds them together, but yet not in absolute contact? This seems to be no strained idea; for expansion and contraction by heat and cold compel us to assume that the particles of matter in general do not touch each other; that there are unfilled spaces between them. In such crystals as we have described these spaces may be considered as alternating with the plates which compose the crystal. From this point of view it seems very natural that the magnetic laminæ should set themselves axial, and the diamagnetic equatorial.

"Our fundamental idea is, that crystals of one cleavage are made up of plates indefinitely thin, separated by spaces indefinitely narrow. If, however, we suppose two cleavages existing at right angles to each other, then we must relinquish the notion of plates and substitute that of little parallel bars; for the plates are divided into such by the second cleavage. If we further suppose these bars to be intersected by a cleavage at right angles to their length, then the component crystals will be little cubes, as in the case of rock-salt and others. By thus increasing the cleavages, the original plates may be subdivided indefinitely, the shape of the little component crystal bearing special relation to the position of the planes. It is an inference which follows immediately from our way of viewing the subject, that if the crystal have several planes of cleavage, but all parallel to the same straight line, this line, in the case of magnetic crystals, will stand axial; in the case of diamagnetic, equatorial. It also follows that in the so-called regular crystals, in rock-salt, for instance, the cleavages annul each other, and consequently, no directive power will be exhibited, which is actually the case."

The above quotation from TYNDALL'S work clearly shows how closely allied are the different forces of crystallization in different directions (which forces determine the planes of cleavage) with the magnetic behaviour of the crystallized medium and lead us to suspect that the forces of cohesion are probably of magnetic nature. The fine points are so completely explained by the magnetic deportment that it is difficult to dissociate the crystalline forces from a magnetic origin. If we assume that these forces are of an electrostatic nature, then it must be admitted that the electrostatic axis of the molecule must coincide with the magnetic axis if the action of a magnetic field is to be decisive, as TYNDALL proved it to be, in isolating the planes of cleavage. But if the electrostatic and magnetic symmetries of the molecules are coincident the application of a field of either nature should induce a double refraction of the same kind in a given liquid. This, however, is not true experimentally, the electric induced double refraction in liquid carbon bisulphide being opposite in sign to the magnetic induced double refraction.* Moreover, in crystalline media, the greatest axes of the ellipsoids representing the magnetic and electric properties of the molecule do not in general coincide. We may therefore say that the evidence points to the conclusion that the forcive which holds the molecules together in a crystalline space lattice is magnetic in nature and not electrostatic.†

DRUDE,‡ in his experiments on the relation between valency and dispersion, suggested that the electron couples constituting the molecule were of two kinds: (1) those of the atoms themselves, the sum of which presumably determine the atomic weight; (2) those of valency which alone are sufficiently free to vibrate synchronously with light waves and hence are particularly concerned in the refraction and dispersion of light.

---

* COTTON and MOUTON, 'Comptes Rendus,' vol. 155, p. 1232, December, 1912.

† [*Note added April 26, 1919.*—After the present communication had passed out of my hands, an important paper "On the Origin of Spectral Series" was published by Sir J. J. THOMSON ('Phil. Mag.,' April, 1919). In this a new theory of atomic structure is suggested in which the atomic nucleus and the revolving electrons play similar rôles to those described on p. 274. Within the contour of the atom, according to Prof. THOMSON, the electrostatic forcive due to the nucleus is of a periodic character and determines a series of spherical or approximately spherical surfaces where the electric force vanishes and over which the periodic motion of the boundary electrons is determined solely by the magnetic field of the atom. This magnetic field is supposed to be radial. If this is the case, these intra-atomic fields must be of the order of magnitude 10^8 gauss $\left(\text{as a simple calculation shows, since } \nu = \dfrac{H\epsilon}{2\pi m}\right)$ to account for the frequencies of the visible spectrum. Still larger intra-atomic fields will exist nearer to the nucleus, of the order 10^9 gauss. These will be sufficient to account for the frequencies of the K series. The infra-red series will be accounted for by fields of the order 10^7 gauss. But this latter value is of the order of the intermolecular magnetic field which has been deduced independently in various ways in the present researches. Moreover, it is to this local field that we have ascribed the rigidity and other properties of crystalline media in general. The frequencies of the infra-red series will, on this view, correspond with the elastic vibrations of the rigid medium in conformity with the quantum theory of specific heats of EINSTEIN and DEBYE as already stated (see Part III., p. 94, and *supra*, p. 259). Reasons have already been given for assigning a magnetic nature to the intermolecular field in crystalline media (see Part III., pp. 101–3, and *supra*, pp. 270–276). This intermolecular magnetic field, which is of the order 10^7 gauss, is suggestive in connexion with Prof. THOMSON's theory, referred to above. On p. 274 (footnote) it was suggested that the forces determining crystalline cohesion are magnetic in nature, the symmetry of the magnetic forces being determined, however, by the electrostatic action of the nucleus. Therefore, in this fundamental sense, the present theory and that of Sir J. J. THOMSON are identical.]

‡ 'Ann. der Phys.,' vol. 14, p. 677 and p. 936, 1904.

The valency or boundary electrons are vibrating under the control of the nucleus, but are less firmly held in the system than those near the nucleus. These valency electrons which have periods corresponding with luminous vibrations are affected by an external magnetic field in accordance with the well-known Zeeman and diamagnetic effects. The highly constitutive nature of the magnetic susceptibility is consistent with this view that the origin of the magnetic property is partly located near the molecular boundary. The nuclear electrons in the free atom will determine symmetry of the molecule, and are directly responsible, by their magnetic effect, for the symmetry of the crystalline grouping. This latter will therefore be determined by the nucleus, which controls the nuclear and boundary electrons, in an indirect manner and the distribution of atomic nuclei in accordance with crystalline symmetry as disclosed by X-ray methods is apparent.

W. H. and W. L. BRAGG have shown the difficulty, even in simple cases, of defining the molecular boundaries in a crystalline space lattice, although in some cases this is possible. But to determine by the X-ray method whether in any given crystal any atom has a special relation to a neighbouring atom would be practically impossible. The X-ray effects which they investigate are determined only by the nucleus or core of the atom and the outer electrons of the atom which contribute to its magnetic property, though they are controlled by the nucleus, are probably distorted by the influences of neighbouring magnetic elements. This distortion, which explains a large number of observed phenomena, defines the molecular boundary within the space lattice and determines a definite chemical molecule. These outer regions of the atom or molecule remain undetected by the X-ray experiments.*

---

* A. E. OXLEY, 'Nature,' No. 4, 1915. The core or electrostatic part of the atom is at a much greater distance from the atomic boundary than are the circular currents which give rise to the magnetic properties. As the intensity of the magnetic field due to a circular current varies inversely as the cube of the distance, and as in a crystalline structure two such circuits may approach so as almost to touch, each electron describing a *small* circle, the local magnetic forcive may be sufficiently large to account for the facts.

The view that the cohesive force in crystalline media is of a magnetic nature was expressed in Part II. of this research, pp. 83–86. It was there stated that in a diamagnetic crystalline medium the molecules are held together by the local magnetic forces due to the revolving electrons. It is possible that each electron is completely bound to its own nucleus by a narrow tube of force, when the molecules would be capable of attracting or repelling one another electromagnetically according to their directions of rotation. The advantage of an electromagnetic cohesive force lies in the fact that by it we can readily see how similar molecules will cling together. Electrostatically such attraction implies an electron transfer which, we know, does not always take place. Reasons are given on p. 277 that the atomic forces which determine the structure of the molecule are in part at least of a magnetic nature. The advantage of

The theory propounded by TYNDALL was called by him the "theory of reciprocal induction," and the direction within the crystalline medium where the molecules had the closest proximity and along which the greatest energy was displayed, he called the "line of elective polarity." This theory is identical with our hypothesis of mutual molecular distortion enunciated at the beginning of Part I. and subsequently confirmed in a variety of ways by other physical phenomena. The direction of closest approach of the molecules, *i.e.*, the line of elective polarity is the line along which the crystal shows the maximum elastic properties. TYNDALL's explanations of the phenomena he had discovered were prophetic. The diamagnetic forces were known to be so minute that the theory of reciprocal induction appeared incredible, and, as a correspondence between Lord KELVIN and TYNDALL shows,* the former expressed emphatically his view that this theory was quite incapable of accounting for the effects observed. On our modern conception of the magnetic structure of matter, this doubt is dispelled and the smallness of the diamagnetic property is no barrier to the theory of reciprocal induction. The effect of applying pressure to a diamagnetic medium, produced, in the direction of the pressure, an increase in the diamagnetic property. This was attributed by TYNDALL to the mutual actions of the diamagnetic polarities which are so minute that their effects, as then understood, would be of such a small order of magnitude that they could not be detected by experiment. On our view of a diamagnetic molecule, which maintains that such a molecule is paramagnetic locally, the effects observed by TYNDALL can be accounted for quantitatively for the local molecular forcives are comparable with those in para- and ferro-magnetic media.

But even in the case of ferro-magnetic media it is not obvious that the magnetic forces are sufficient to explain the mechanical phenomena unless we realise the localised nature of the forcive. If we take, for example, a crevasse of the usual

---

assigning a magnetic nature to the forces of valency is clear, for in this way, without admitting an electron transfer between the various atoms forming the molecule, we can secure the necessary attraction, and this by a fixed or directed force which at the same time is compatible with a characteristic orbital frequency such as appears to be necessary to account for ordinary absorption, magnetic rotation, and diamagnetic phenomena. The possibility of a satisfactory interpretation of many problems suggested by stereochemistry, in terms of the magnetic force due to revolving electrons, has been ably expounded by A. L. PARSON ('Smithsonian Miscellaneous Collections,' vol. 65, No. 11, a paper to which, on account of war service, I have only recently had access). Though PARSON's theory involves new difficulties in connexion with the distribution of positive electricity in the atom, the advantages from a chemical standpoint which he secures by the introduction of magnetic forces of chemical combination cannot be denied. Granting this, it is natural to suppose that the cohesive forces, which hold the molecules in position in a space lattice, will be residual magnetic forces, and that they will closely resemble, in distribution at least, the atomic forces determining the configuration of the molecule. It will be of great interest to see how far such magnetic cohesive forces are capable of interpreting the spacing of molecules in a crystalline lattice in accordance with the distribution disclosed by X-ray analysis. Magnecrystallic action, as we have seen, is explicable in this way.

* Various letters, 'On Diamagnetism and Magnecrystallic Action,' 1870.

conventional dimensions, within an iron crystal where the saturation intensity is I, the mechanical stress is $2\pi I^2$, which is far smaller than the ultimate tensile strength of the iron. Indeed, EWING* remarks "we may, if we please, regard the magnetic molecules as pulling at one another across any imaginary interface, while the stress with which they pull is balanced by thrust in the framework of the iron, but neither the pull nor the thrust is competent to explain the mechanical strains." The above value of the stress, viz., $2\pi I^2$ is obtained by taking a crevasse whose gap, although "physically small," is sufficiently wide to accommodate several molecules in line. If we take a narrower crevasse, approaching "mathematical smallness" in width of gap, we obtain a measure of the forcive acting between the molecules, and this includes the localised forcive NI or $\alpha'_c I$ which is of the order 10^7 gauss. The localised stress across this interface is $\frac{1}{2}NI^2$ or $\frac{1}{2}\alpha'_c I^2$, which we have seen to be of the order 2×10^9 dynes per square centimetre. This is of the same order as the ultimate tensile strength of crystalline media both ferro-magnetic and diamagnetic.

The magnetic resistance of joints is interesting in connexion with the localised nature of the molecular field in iron. It has been shown by Sir J. J. THOMSON and H. F. NEWALL† that the susceptibility of an iron bar is much reduced if it is severed and the two parts put in contact. Later, Sir JAMES EWING and W. LOW‡ investigated this effect in a more exhaustive manner when the joints were carefully trued up and also for rough joints, under varying pressures. They found that for a carefully planed joint a compressive stress of 226 kilogrammes per square centimetre restored almost completely the loss of magnetic property produced by cutting, but that this stress had only a small restorative effect in the case of a rough joint. In the latter case, we may suppose that the number of points of contact between the two parts of the bar is small, in the former that the two portions are in contact over a large percentage of the available area of contact. Under a compressive stress of 226 kilogrammes per square centimetre, it appears that in the trued-up specimens the order of contact of the molecules is the same as in the uncut metal and therefore this stress is a measure of the internal stress within the material. As 226 kilogrammes per square centimetre is equal to $0 \cdot 5 \times 10^9$ dynes per square centimetre, this stress, although lower, is comparable with that calculated on p. 252, and we may regard the width of the resulting crevasse as approaching mathematical smallness, the spheres of influence of the molecules on either side of it overlapping to an extent comparable with the overlap in the interior of the uncut bar (see also Part III., p. 89). But even with the most carefully faced junction there will be irregularities, coarse compared with molecular dimensions, and in such regions the localised nature of the molecular field will determine a finite air gap which would account for the difference of stress mentioned above. Perfectly faced surfaces of soft iron or mild steel (annealed) might be expected, under the influence of an external magnetic field, to form a perfect junction, in other words to become welded together.

---

* 'Magnetic Induction in Iron and other Metals,' p. 254.

† 'Proc. Camb. Phil. Soc.,' 1887.

‡ 'Phil. Mag.,' September, 1888.

TAYLOR JONES* has obtained an induction as high as 74,200 Maxwells in soft iron under strong fields. The tension necessary to pull the surfaces apart in this case will be $\frac{B^2}{8\pi} = \frac{7{\cdot}4^2 \times 10^8}{8\pi}$, or 2×10^8 dynes/square centimetre, which is about a twenty-fifth of the tensile strength of the material.

(6) ON A MAGNETIC THEORY OF CHEMICAL COMBINATION.

On the theory of chemical action developed by Sir J. J. THOMSON,† the determining feature of an atom from the point of view of chemical combination is the number of positive valency electrons it possesses. These electrons are dragged from their loose attachment to the nucleus, during chemical combination, and pass from one atom to another. The two originally neutral atoms thus become oppositely charged and so attract one another and form, as it were, an electric doublet. Let us look at this problem from the magnetic standpoint. Each electron orbit is equivalent to a small magnetic doublet and it is interesting to enquire how far the magnetic forces of such doublets may represent the force of chemical affinity. Recent work on radio-activity, the wide deflections of β-rays, and the diffraction of X-rays, all point to a localisation of the electrostatic charges in a minute core or nucleus. Round this nucleus, and under its control, the valency electrons (in part responsible for the magnetic properties) rotate. It is conceivable, therefore, that the magnetic forces, in addition to the important role they play in crystallization, may also in part be responsible for the forces of chemical affinity.‡

If, during chemical combination, there is a definite transfer of valency electrons from one atom to another, we should expect to find an abrupt change in the magnetic behaviour of an atom before and after chemical combination. If, on the other hand, there is no such electron transference, we might expect that the atoms would preserve their magnetic properties, which would be more or less of an additive nature. In a remarkable series of investigations, PASCAL§ has shown that in a very large number of organic compounds, the molecular susceptibility is the sum of the atomic susceptibilities of the component atoms, provided the molecule contains no peculiarity of molecular configuration—such, for instance, as the ethylene linkage, unsaturated atom, or complex nucleus. Thus if χ_M is the molecular susceptibility, and χ_A the atomic susceptibility of a component atom, we have

$$\chi_M = \Sigma\chi_A + \lambda$$

where the summation extends to all the atoms in the molecule and λ is a positive or negative constant for a certain type of peculiarity of molecular configuration. In normally saturated compounds $\lambda = 0$. The elements carbon, hydrogen, chlorine, bromine and iodine have constant atomic susceptibilities in a large variety of simple and complex organic compounds. This suggests that the origin of the valencies of these elements is also the origin of a definite amount of diamagnetism, under the influence of a definite magnetic field. In other words, a hydrogen atom, in whatever organic compound it is found, has a constant atomic susceptibility equal to $-30{\cdot}5 \times 10^{-7}$, while the carbon atom has a constant atomic susceptibility equal to $-62{\cdot}5 \times 10^{-7}$, and so on.

---

* 'Phil. Mag.,' vol. xli., p. 165, 1896.

† 'The Corpuscular Theory of Matter,' 1907, Chap. VI.

‡ See footnote p. 274. Also W. M. HICKS, 'Roy. Soc. Proc.,' A, vol. 90, p. 356.

§ 'Ann. de Chim. et de Physique,' sér. 8, vol. 19, p. 5, 1910.

This result of PASCAL'S, in conjunction with—

(1) The enormous magnitude of the local molecular field in diamagnetic media, and

(2) The conception of diamagnetism as due to an induction effect in oppositely spinning electrons (as developed in Parts I., II. and III.), led me to suspect that the magneton may be a constituent of the diamagnetic hydrogen molecule. The calculation showed* that if there is one electron in each hydrogen atom whose period is equal to

---

* 'Roy. Soc. Proc.,' A, vol. 95, p. 58, 1918.

At the time this paper was written, I was out of touch with the latest available data concerning the values of AVAGADRO'S constant (N) and the ratio ϵ/m. The calculation was to determine M from the relation

$$M = -\frac{\chi \,.\, 4\pi m}{N \,.\, n \,.\, \epsilon \,.\, \tau}.$$

Taking χ = molecular susceptibility of hydrogen,
$= -61{\cdot}0 \times 10^{-7}$ (PASCAL),
$N = 6{\cdot}06 \times 10^{23}$ (MILLIKAN),
$\epsilon/m = 1{\cdot}77 \times 10^{7}$ e.m.u. (BUCHERER),
$n = 2$, the number of electrons per molecule,
$\tau = 2{\cdot}19 \times 10^{-15}$ sec., the period of revolution for the line Hα,

we find on calculation

$$M = 16{\cdot}3 \times 10^{-22} \text{ for the moment of the magneton.}$$

This gives for r, as calculated from $M = \frac{\pi\epsilon r^2}{\tau}$, the value $0{\cdot}85 \times 10^{-8}$ cm.

In this connexion it should be pointed out that, on PLANCK'S theory of quanta of energy, the constant h is consistent with the existence of a unit of magnetism. Assuming, as NICHOLSON and BOHR have done, that the angular momentum of the electron is an integral multiple of $\frac{h}{2\pi}$, CHALMERS showed that the magnetic moment of the electron orbit is

$$M = \frac{\epsilon}{m} \,.\, \frac{h}{4\pi}.$$

This gives $M = 92{\cdot}4 \times 10^{-22}$ e.m.u., which is 5 times the experimental value of the moment of the magneton. If we leave aside PLANCK'S theory of energy quanta and adopt instead SOMMERFELD'S theory of quanta of action, LANGEVIN showed that a remarkable relation between h and M exists. He found

$$M = \frac{\epsilon}{m} \,.\, \frac{h}{24\pi},$$

when the law of attraction between the nucleus and the electron is the inverse square. This gives for the magnetic moment of the electron orbit $M = 15{\cdot}4 \times 10^{-22}$ e.m.u., a value nearly equal to the most recent experimental value of the moment of the magneton, viz., $18{\cdot}5 \times 10^{-22}$ e.m.u.

that of the line H_α, then the molecular diamagnetic susceptibility of hydrogen $(-61 \cdot 0 \times 10^{-7})$ can be accounted for, and each electron orbit of radius 10^{-8} cm. has a magnetic moment $+16 \cdot 3 \times 10^{-22}$, nearly equal to that of the magneton, $+18 \cdot 5 \times 10^{-22}$.

The atomic susceptibility of carbon in combination is shown by PASCAL to be $-62 \cdot 5 \times 10^{-7}$, and in connexion with the additive law this value is consistent with the experimental values of the molecular susceptibilities. The mean experimental value of the atomic susceptibility of diamond is $-59 \cdot 0 \times 10^{-7}$.* The mean of these values is $-60 \cdot 7 \times 10^{-7}$, which is probably as accurate a value as is available at present. But this value is almost exactly twice that of the atomic susceptibility of hydrogen $-30 \cdot 5 \times 10^{-7}$. Probably therefore the atom of carbon contains two magnetons. As to the period of revolution we have

$$M = \tfrac{1}{2} \cdot \epsilon \cdot \omega r^2 = \frac{\pi \epsilon r^2}{\tau}$$

where

M = moment of orbit,
ϵ = electron charge in e.m.u.,
r = radius of orbit,
τ = period,

and this implies that the period τ for the carbon atom is not equal to that of the line H_α unless $r = 10^{-8}$ for the carbon atom.

The sum of the atomic susceptibilities of the atoms in the group CH_2 is

$$-123 \cdot 5 \times 10^{-7}.$$

Experimentally, PASCAL showed that the difference of molecular susceptibility in a number of compounds whose constitution differed by this group was

$$-118 \cdot 7 \times 10^{-7}.$$

The mean of these two values is $-121 \cdot 1 \times 10^{-7}$ which is probably very near the true value. But this is almost exactly four times the value of the atomic susceptibility of hydrogen, viz., $-30 \cdot 5 \times 10^{-7}$, and in this combination, CH_2, we may say that each hydrogen atom has one magneton and the carbon atom has two magnetons.

An ethylene linkage, according to PASCAL, lowers the diamagnetic molecular susceptibility of the compound by $+57 \times 10^{-7}$, while two or more such linkages lower it by $+110 \times 10^{-7}$. These values are respectively nearly equal to twice and four times the atomic susceptibility of hydrogen.

As the additive law holds in the case of the simpler liquid hydrocarbons, it will hold for all the others which differ only by CH_2 groups, and therefore, taking the values of λ into account, any member of the three homologous series

$$C_n \cdot H_{2n+2}, \qquad C_n \cdot H_{2n}, \qquad C_n \cdot H_{2n-2}$$

---

* HONDA, 'Ann. der Phys.,' vol. 32, p. 1044, 1910, gives $-59 \cdot 5 \times 10^{-7}$; OWEN, 'Ann. der Phys.,' vol. 37, p. 693, 1912, gives $-58 \cdot 5 \times 10^{-7}$.

may be considered as containing respectively

$$4n+2, \quad 4n-2, \quad 4n-6 \text{ magnetons.}$$

Moreover, PASCAL has shown that the magnitude of χ_{CH_2} is quite independent of the presence of nitrogen, oxygen, or sulphur in the compound, so that, where the additive law holds for such complex molecules, it will hold for the whole series of *additive* compounds formed with the group.

PASCAL also investigated the halogens, fluorine, chlorine, bromine and iodine. The atomic susceptibilities of these elements, as deduced from the molecular susceptibilities of organic compounds in which they are contained, are given in column 2, the approximate number of magnetons per atom is given in column 3.

Substance.	$\chi_A . 10^7$.	n.
Fluorine	$-65 \cdot 5$	2
Chlorine	$-209 \cdot 5$	7
Bromine	$-319 \cdot 2$	10
Iodine	$-465 \cdot 0$	15

It must be admitted, however, that these last four results do not warrant an extension of the magneton theory to these elements. Moreover, although the results for the hydrocarbons mentioned above are very suggestive, yet there remain difficulties, such for example as the values of λ, which for the benzene ring is equal to -15×10^{-7}, the interpretation of which does not fall into line with the magneton view. Further difficulties are met with in the cases of nitrogen and oxygen.

Perhaps these difficulties should be expected, since it has been proved that the additive law breaks down for many atoms, especially as regards the metallic elements. When, in addition, we take into consideration that the mutual disturbances of the electron orbits, in atoms containing a relatively large number of electrons, have been neglected, the agreement is probably as good as could be expected. Thus in the case of the hydrocarbons considered above, compounds which show no electrolytic dissociation and in the formation of which no transfer of electrons takes place from one atom to another on combination, we might expect that the addition of a hydrogen atom or a CH_2 group would add a definite amount of diamagnetism to the compound. But in the case of the metals and some other elements, chemical combination may be accompanied by the transference of electrons, *i.e.*, by a breakdown of the magnetic elements of the atoms. In such cases the additive law could not hold.

Thus iron-carbonyl (Fe $(CO)_5$) and nickel-carbonyl (Ni $(CO)_4$) are diamagnetic[*]; potassium ferri-cyanide is paramagnetic, while potassium ferro-cyanide is diamagnetic.[†] It would appear that in these cases the loss of magnetic property of the iron and nickel atoms is due to a transfer of valency electrons, *i.e.*, it involves an electric charging up of the atoms. The behaviour of the oxygen atom in organic compounds, in compounds with chlorine and in metallic oxides, where it acts always as an electronegative element, may possibly be accounted for in the same way. Free oxygen and ozone are strongly paramagnetic, but no semblance of an additive nature of the magnetic property is found in any of the oxygen compounds.

The appearance of strong magnetism in the Heusler alloys and its disappearance in manganese steels, are similar effects, dependent on the formation of intermetallic chemical compounds accompanied by an electron transfer.

There are many paramagnetic substances which possess molecular magnetic moments comparable with, and in some cases much superior to, those shown by ferromagnetic substances. The apparent feeble susceptibility they possess is due to the fact that, with the largest magnetic field which can be applied, we can never produce anything like a saturation effect. In fact, according to WEISS, the molecule of cobalt chloride, $Co.Cl_2$ contains 25 magnetons, while an atom of cobalt below the critical temperature contains 9 magnetons. We may well ask ourselves—what is the nature of the process by which the addition of a diamagnetic substance, H.Cl to cobalt, produces such a large increase in the number of magnetons per molecule.[‡] Assuming the work of WEISS holds good, and there is certainly a very considerable amount of evidence in favour of his theory, we can interpret this result either by supposing that the atom of cobalt really contains more than 9 magnetons,[§] or else that the *diamagnetic* acid supplies the additional magnetons when it acts on the cobalt to form the chloride. In either case, it seems that we must admit that a molecule may possess systems of magnetons which, in certain circumstances, are so arranged to counterbalance one another, producing no additional moment of the molecule.[||] These magnetons would contribute nothing to the paramagnetic or ferro-magnetic property of a substance and could not be included in WEISS'S theory. The grouping of these "latent" magnetons, according to our extended view, would be perturbed by the union of the cobalt atom with the Cl ion, in a manner similar to that by which an external field reveals the spontaneous magnetization in iron (as interpreted on EWING'S theory) by orientating groups which were formerly so constituted as to show no magnetic effect externally. If this is so, then we are only justified in assuming that purely diamagnetic molecules contain groups of magnetons so arranged that the molecule possesses no resultant magnetic moment. For example, while iron below the critical temperature possesses, according to WEISS, 11 magnetons to the atom, the molecule of ferric sulphate possesses 30, ferric chloride 28, sodium ferro-pyrophosphate 26, sodium ferrous oxalate 27.* Again, nickel below the critical temperature possesses 3 magnetons to the atom, above the critical temperature either 8 or 9 magnetons to the atom, while a molecule of nickel sulphate contains 16 magnetons. In general the number of magnetons, per molecule of a salt of a ferro-magnetic element, is large compared with the number of magnetons associated with an atom of the pure ferro-magnetic. The fact that cupric salts are paramagnetic while cuprous ones are diamagnetic is interesting from our point of view. Although copper is diamagnetic, yet a molecule of cupric sulphate contains 10 magnetons. It seems as if the large local atomic fields, which have been recognised in diamagnetic and ferro-magnetic molecules, have the power, when the molecules approach so that their fields overlap, to upset the equilibrium of the atoms in combination and redistribute their magnetic elements. This is easily possible when the great intensity of the local molecular field is borne in mind. In most cases a diamagnetic molecule, on account of its symmetry, would, under the influence of such a field, remain diamagnetic, but each orbit would be distorted by the field and the susceptibility of the substance would be slightly modified.

---

[*] A. E. OXLEY, 'Proc. Camb. Phil. Soc.,' vol. 16, p. 102, 1911.

[†] J. S. TOWNSEND, 'Phil. Trans. Roy. Soc.,' A, vol. 187, p. 547, 1896.

[‡] The specific susceptibility of cobalt chloride is 90×10^{-6}, that of hydrochloric acid $-0 \cdot 80 \times 10^{-6}$.

[§] Some of which are self-compensated.

[||] *I.e.*, of the molecule as a whole.

(7) ON SOME ANOMALIES IN THE MAGNETIC ROTATION EFFECT.

Diamagnetic media are in general dextro-gyric. The only exception is titanium chloride which is lævo-gyric.† Paramagnetic media are sometimes dextro-gyric and sometimes lævo-gyric, while the ferro-magnetic elements, iron, nickel and cobalt are all dextro-gyric. At present no theory seems capable of accounting for these anomalies and it is therefore interesting to examine to what extent the local molecular field may cause the effects observed. At one time it was thought that the direction of rotation probably depended merely on the diamagnetic or paramagnetic property of the molecule, but experiment soon disproved this generalisation. VOIGT ‡ suggested that the production of an intense reverse field, when the external field was applied, would account for the effects, but no physical explanation of a possible origin for this intense reverse field was given.

On the views of diamagnetic and paramagnetic polarity developed in these researches, the necessary fields demanded by VOIGT are found in the immediate neighbourhood of the molecular boundary. In diamagnetic liquids the molecules have zero magnetic moments and their axes are distributed at random. The application of an external field produces a minute diamagnetic polarity, which differential effect determines the magnitude and sign of the rotation. When a polarised ray passes over such molecules, it will be rotated considerably locally, but only to be rotated in the opposite direction to an almost equal extent in a neighbouring molecule or in another part of the same molecule.

---

* 'Journal de Phys.,' vol. 1, sér. v., p. 974, 1911.

† For solutions of salts in water it should be noted that VERDET'S constant for the solute is to be regarded as negative if VERDET'S constant for the solvent is less than 0·0130, which is the value of the constant for pure water at 20° C.

‡ See P. ZEEMAN, 'Researches in Magneto Optics,' p. 185.

This will also be true if the diamagnetic medium is crystalline. Hence the final rotation will be a relatively small differential effect compatible with the diamagnetic effect of the medium in bulk. We shall omit the exceptional case of titanium chloride for the present and pass on to consider the ferro-magnetic media iron, nickel and cobalt.

In ferro-magnetic media, on account of the continuity of magnetic induction, there is an enormous reverse local field, and if this acts over regions of the molecule containing magnetically active electrons, a large rotation will be produced, which will not be compensated in neighbouring molecules, when the ferro-magnetic material is saturated. Hence we should expect a very large rotation to be produced by such media. This has been confirmed experimentally, in the cases of iron, nickel and cobalt, by KUNDT.

Paramagnetic solutions lie in an intermediate category. The application of an external magnetic field causes a certain amount of molecular orientation depending on the temperature. Such orientation causes a reverse field over a part of the system, but over the neighbouring molecules in combination with the one considered, the field depends upon the difference between the reverse local field and the applied field, which difference may be positive or negative. If, in such regions, there are electrons capable of orientating the polarised beam, the rotation may be dextro- or lævo-gyric, according as the resultant field is opposite to, or in the same direction as, the applied field. This will not explain the exceptional case of titanium chloride unless this molecule possesses some peculiar dissymmetry, whereby, in spite of its diamagnetic nature, it becomes orientated under the applied field. Such orientation is the basis of the magnetic double refraction theory developed by LANGEVIN* and confirmed experimentally in some respects by COTTON and MOUTON† (see Part III., p. 87). An alternative explanation of the behaviour of the titanium chloride molecule may be found in a rotation of the paramagnetic titanium atom relative to the compound molecule.

BECQUEREL has deduced an interesting relation‡ connecting the magnetic rotation (r) with the Zeeman coefficient $\left(Z = \frac{\epsilon}{4\pi m}\right)$, the applied field H, wave-length λ and refractive index μ, viz.,

$$r = -2\pi \cdot \frac{ZH\lambda}{c} \cdot \frac{\partial\mu}{\partial\lambda}, \qquad (1)$$

where c is the velocity of light.

---

* LANGEVIN, 'Le Radium,' vol. 7, p. 251, 1910.

† 'Ann. de Chim. et de Phys.,' sér. viii., vol. 19, p. 155, 1910.

‡ SCHUSTER, 'Theory of Optics,' p. 307.

We shall assume that this relation applies to all media. On our theory H will be the sum of the applied and local fields and will be considered positive if it is in the direction of the external field. Z we shall take to be negative, *i.e.*, the negative electron is always responsible for the Zeeman effect.

In diamagnetic media, in general, H is positive, $\frac{\partial\mu}{\partial\lambda}$ is negative, and therefore r is negative, or clockwise looked at along the direction of H. The rotation is dextro-gyric.

In ferro-magnetic media, H is the resultant of the applied field and a very large reverse molecular field, so that H is very large and negative; $\frac{\partial\mu}{\partial\lambda}$ is positive. Hence again r is negative and the rotation is dextro-gyric as in diamagnetic media.

In paramagnetic solutions, H may be positive or negative over magnetically active atoms, and therefore the sign of r may be positive or negative. Hence some paramagnetic solutions will be dextro- and some lævo-gyric.

In the ferro-magnetic elements the magnitude of the rotation is remarkably high. Thin films of saturated iron show a rotation of the order 260 million times that of carbon bi-sulphide subjected to an external field of one gauss. To test this we may write

$$r_{\text{Fe (saturated)}} = -2\pi \, . \, \frac{Z_{\text{Fe}} \, . \, H_c}{c} \, . \, \lambda \, . \, \frac{\partial\mu_{\text{Fe}}}{\partial\lambda} \text{ for iron,}$$

$$r_{\text{CS}_2\,(\text{H}=1\text{ gauss})} = -2\pi \, . \, \frac{Z_{\text{CS}_2} \, . \, H}{c} \, . \, \lambda \, . \, \frac{\partial\mu_{\text{CS}_2}}{\partial\lambda} \text{ for carbon bisulphide.}$$

where H_c is the reversed molecular field, of the order $-6{\cdot}5\times10^6$ gauss, $H = 1$ gauss. The Zeeman coefficients Z_{Fe} and Z_{CS_2} we shall take to be of the same order, hence :—

$$\frac{r_{\text{Fe (saturated)}}}{r_{\text{CS}_2\,(\text{H}=1\text{ gauss})}} = +\frac{H_c \, . \, \frac{\partial\mu_{\text{Fe}}}{\partial\lambda}}{H \, . \, \frac{\partial\mu_{\text{CS}_2}}{\partial\lambda}} \quad \ldots\ldots\ldots \quad (2)$$

For iron*

$\lambda\times10^8$.	μ.
2570	1·01
4410	1·28
5890 (D)	1·51

Near

$$\lambda = 5890\times10^{-8}, \qquad \partial\mu_{\text{Fe}}/\partial\lambda = +\frac{0{\cdot}23}{1480\times10^{-8}}.$$

---

[*] 'Smithsonian Physical Tables,' p. 196.

For carbon bisulphide*

$\lambda \times 10^8$.	μ.
4340 (H_γ)	1·6920
4860 (H_β)	1·6688
5890 (D)	1·6443

Near

$$\lambda = 5890 \times 10^{-8}, \qquad \partial\mu_{CS_2}/\partial\lambda = -\frac{0{\cdot}0255}{1030 \times 10^{-8}}.$$

From (2)

$$\frac{r_{\text{Fe (saturated)}}}{r_{CS_2\text{ (H = 1 gauss)}}} = +\frac{6{\cdot}53 \times 10^6 \times 0{\cdot}23 \times 1030 \times 10^{-8}}{0{\cdot}0255 \times 1480 \times 10^{-8}} \doteqdot 40 \times 10^6.$$

For nickel†

$\lambda \times 10^8$.	μ.
2750	1·09
4410	1·16
5890 (D)	1·30

Near

$$\lambda = 5890 \times 10^{-8}, \qquad \frac{\partial\mu_{Ni}}{\partial\lambda} = +\frac{0{\cdot}14}{1480 \times 10^{-8}},$$

and

$$\frac{r_{\text{Ni (saturated)}}}{r_{CS_2\text{ (H = 1 gauss)}}} = +\frac{6{\cdot}35 \times 10^6 \times 0{\cdot}14 \times 1030 \times 10^{-8}}{0{\cdot}0255 \times 1480 \times 10^{-8}} \doteqdot 25 \times 10^6.$$

For cobalt‡

$\lambda \times 10^8$.	μ.
2750	1·41
5000	1·93
6500	2·35

Near

$$\lambda = 5890 \times 10^{-8}, \qquad \frac{\partial\mu_{Co}}{\partial\lambda} = +\frac{0{\cdot}42}{1500 \times 10^{-8}}.$$

Taking the reverse molecular field equal to that in iron

$$\frac{r_{\text{Co (saturated)}}}{r_{CS_2\text{ (H = 1 gauss)}}} = +\frac{6{\cdot}53 \times 10^6 \times 0{\cdot}42 \times 1030 \times 10^{-8}}{0{\cdot}0255 \times 1500 \times 10^{-8}} \doteqdot 87 \times 10^6.$$

These values, although rather small, are comparable with the enormous ratios of the rotations of saturated ferro-magnetics to diamagnetics, as actually observed experimentally, and point to the importance of the local molecular forcive in magneto-optic phenomena. It should be noted that the value of the reverse field we have taken is that of the *intermolecular* field. Within the atom the reverse field will probably be greater than $-6{\cdot}5 \times 10^6$ gauss. If it is of the order -5×10^7 gauss, the above ratios become 320×10^6 for iron, 200×10^6 for nickel and 700×10^6 for cobalt. Intra-atomic fields of the order 10^8 gauss are required by HUMPHRIES* to explain the pressure shift of spectral lines and by RITZ† in his theory of spectral series (see also Part III., p. 100, and *supra*, p. 273, footnote).

---

* *Loc. cit.*, p. 192.

† *Loc. cit.*, p. 196.

‡ *Loc. cit.*, p. 195.

(8) SUMMARY OF CONCLUSIONS.

(I.) The applications of the local molecular forcive, in diamagnetic, paramagnetic and ferro-magnetic media, have in the present research been extended to interpret the ultimate tensile strength of crystalline and vitreous media. It has been shown by EWING and ROSENHAIN that the permanent set which occurs prior to breaking is due to slipping along the cleavage planes *within* the individual crystalline grains. We should therefore expect that the material would be fractured when the applied mechanical stress is equal to that produced internally by the local molecular forcive. The internal stress within the material is shown to be of the order 2×10^9 dynes per square centimetre which is approximately the mean value of the ultimate tensile strengths of crystalline and vitreous media (pp. 250–259).

(II.) As a consequence of this internal stress, the energy per unit volume will be 2×10^9 ergs, and this energy, which is over and above that which exists in the fluid state, should be a measure of the latent heat of fusion per cubic centimetre. This test which was applied in Part III. to test the order of the local forcive, has been extended to a variety of organic and inorganic media, including the metals and is found to accord with the experimental values to the right order (pp. 253–4).

(III.) Since the forces under which the molecules vibrate are those to which we ascribe the elastic properties of crystalline media, the results obtained are consistent with the theory of specific heats developed by DEBYE, in which the specific heat is attributed to purely translational vibration, and it has been shown (Part III.) that, near the fusion point, the rotational energy acquired by the molecules will give a measurable departure from this theory which is actually observed experimentally. As we should expect, it is found that the elastic constants of a variety of ferro-magnetic, diamagnetic and paramagnetic media are of the same order, several diamagnetic and paramagnetic media even surpassing steel in their power to resist distortion (pp. 257–259).

(IV.) Any change of internal pressure will be accompanied by a change of volume defined by the compressibility of the medium and dependent as to sign upon peculiarities of the molecular configuration. It has been shown that the energy change which occurs on crystallization is compatible with a volume change of the same order of magnitude as that accompanying crystallization, and we may therefore interpret the change of volume on crystallization as a magneto-striction effect of the local molecular forcive. The magneto-striction effect depends on molecular orientation which is proportional to the square of the magnetic forcive (pp. 259–263). In Part III. it was shown that the double refraction of crystalline media can be interpreted as due to the magnetic double refraction effect of the local molecular forcive which orientates the molecules into a crystalline space lattice. This effect is also proportional to the square of the magnetic forcive and the two effects mutually support one another.

---

* 'Astrophysical Journal,' vol. 23, p. 232, 1906; vol. 35, p. 268, 1912.

† 'Ann. der Phys.,' vol. 25, p. 660, 1908.

(V.) The above results are interesting in connexion with TAMMANN's theory of the closed region of stability of the crystalline state, as represented on the pressure temperature diagram. TAMMANN's experimental work gives an alternative method of determining intrinsic pressures, but the results are notably higher than those found in other ways. Possibly this is due to extrapolation over a wide pressure range. The pressure temperature diagrams showing the fusion curve are instructive in dealing with problems relating to thermal evolution (or absorption), and volume changes and possible interpretations of these peculiarities in the ferro-magnetic elements, iron, nickel and cobalt, have been given (pp. 263–266).

(VI.) A physical interpretation has been given of the large values of the coefficients, N and α'_c, of the molecular fields in ferro-magnetic and diamagnetic crystalline media respectively. These coefficients are the reciprocals of the limiting local susceptibilities of the media under field strengths equal to the respective molecular fields. The local susceptibility of a diamagnetic molecule is comparable with that of a ferro-magnetic molecule and the two vary in the same way with field strength and temperature. In diamagnetic media, however, magnetic hysteresis will be inappreciable, since the molecule as a whole possesses a zero magnetic moment. Nevertheless, mechanical hysteresis in diamagnetic media will be of the same order as in ferro-magnetic media (pp. 267–270 and p. 257).

(VII.) From TYNDALL's experiments on the deportment of paramagnetic and diamagnetic crystals in a magnetic field, the positions of the planes of cleavage can be traced. These results show that the forces responsible for crystalline symmetry are very probably of a magnetic nature. If the forces are of an electrostatic nature, then, since an electric field must disclose the same planes of cleavage, the electric and magnetic symmetries must coincide. This is not the case however. The magnetic forces are partly due to the valency or boundary electrons whose orbits are controlled by the atomic nuclei. The nuclei determine the crystalline symmetry, indirectly, through the medium of the magnetic forces of the electrons. This conclusion is not at variance with the results of X-ray diffraction experiments; the latter determine only the positions of the diffuse diffracting cores and give no indication as to the outer regions of the atom which determine the valency forcives. Hence, in spite of the disclosure of the X-ray methods, it is maintained that within a crystalline medium the molecules, though distorted, are still essentially integral units and that it is possible to imagine a surface enclosing each (pp. 270–277).

(VIII.) The smallness of the nucleus disclosed by recent work on radio-activity suggests that in addition to the part played by the magnetic forces in crystallization, these forces are in part responsible for chemical combination (footnote, p. 274). The theory of chemical combination developed by Sir J. J. THOMSON implies the transference of electrons from one atom to another whereby the atoms become oppositely charged. Such a transference of the valency electrons implies a complete readjustment of the magnetic property and therefore this property could not be of an additive nature. This is borne out by a large amount of experimental evidence.

If the chemical combination takes place without a transference of the valency electrons from one atom to another, which probably happens in many organic compounds in which electrolytic dissociation does not take place, we might expect the magnetic properties to be more or less of an additive nature. PASCAL'S work confirms this view (pp. 277–281).

(IX.) The author has previously shown that the magneton may be a constituent of the diamagnetic hydrogen molecule. It appears that this idea may be extended to carbon and the hydrocarbons in general, where the molecular susceptibility can be directly calculated from the atomic susceptibilities of the component atoms. The results, however, are not so convincing when extended to other elements, but the fact that departures from the additive law occur in such cases leads us to suppose that some disturbing influence has been neglected. WEISS'S work on salts of the ferromagnetic elements, taken in conjunction with our present conception of diamagnetism, suggests that diamagnetic substances contain magnetons, compensated so as to produce a diamagnetic effect of the medium in bulk. The forces of chemical combination may, however, perturb this state, and by rearranging the magnetic elements give rise to a compound possessing more magnetons than are contained in an unbalanced state in a ferro-magnetic element. Examples of this are given and also of the reverse effect which may equally well arise (pp. 281–282).

(X.) The principle of the continuity of magnetic induction, as applied to the local molecular forcive, suggests a possible interpretation of known anomalies of the magnetic rotational effect. In paramagnetic solutions, the dextro-gyric or lævo-gyric property is attributed to a differential effect of the reversed local field and the applied field over rotationally active electrons, the sign of the effect depending on the direction of the resultant field acting over these electrons.

The very powerful rotational effect of thin films of iron, nickel and cobalt, in comparison with that shown by carbon bisulphide, has also received an interpretation in terms of the local molecular forcive (pp. 282–286).

(XI.) It is considered that the above conclusions when conjoined with those obtained in Parts I., II., III., which bring the theory into line with the magnetic atom fields of **Ritz** and probably suggest an origin of spectral series, amply justify the importance of the magnetic forcive in crystalline and vitreous media. The magnitude of this local magnetic forcive, first calculated to interpret the change of diamagnetic susceptibility observed on crystallization of a large number of organic compounds experimented upon in Parts I. and III., has been found capable of correlating a number of additional physical phenomena of wide difference of origin. It is hoped to continue these applications to other branches of optics including spectral series and optical activity.

(XII.) [*Added February 28, 1920.*—It has been established that the intermolecular field, in all crystalline media, is of the order 10^7 gauss. The electrons, within the free atom, are controlled by electrostatic, and possibly also by magnetic, forces, whose origin lies in the core. When the atoms are grouped into a definite space lattice, the cohesive force between them is of a magnetic nature, and the rigidity of such media is due to the localised mechanical stress, exerted at definite points across the atomic "surface"; the electrons revolving in small circles in adjacent atoms (p. 274). This mechanical stress, due to the local magnetic forces, is responsible for the change of specific susceptibility, and other properties, on crystallization; and is balanced by the stress due to the distortion of the internal electrostatic configuration of the atoms. In this way, a balance is secured between the electrostatic and magnetic stresses; these stresses predominating alternately, as we pass through the crystalline structure, thus giving rise to a system in equilibrium.

The rotation of electrons in *small* circles, at definite points near the atomic "surface," is suggestive in connexion with the theory of directed valencies required to explain stereo-chemical phenomena (p. 274, footnote).]

IX. *A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations made at the Total Eclipse of May 29, 1919.*

By Sir F. W. DYSON, *F.R.S., Astronomer Royal, Prof.* A. S. EDDINGTON, *F.R.S.*, *and Mr.* C. DAVIDSON.

(*Communicated by the Joint Permanent Eclipse Committee.*)

Received October 30,—Read November 6, 1919.

[PLATE 1.]

CONTENTS.

Page
I. Purpose of the Expeditions 291
II. Preparations for the Expeditions 293
III. The Expedition to Sobral 296
IV. The Expedition to Principe 312
V. General Conclusions 330

I. PURPOSE OF THE EXPEDITIONS.

1. THE purpose of the expeditions was to determine what effect, if any, is produced by a gravitational field on the path of a ray of light traversing it. Apart from possible surprises, there appeared to be three alternatives, which it was especially desired to discriminate between—

(1) The path is uninfluenced by gravitation.

(2) The energy or mass of light is subject to gravitation in the same way as ordinary matter. If the law of gravitation is strictly the Newtonian law, this leads to an apparent displacement of a star close to the sun's limb amounting to $0''\cdot87$ outwards.

(3) The course of a ray of light is in accordance with EINSTEIN's generalised relativity theory. This leads to an apparent displacement of a star at the limb amounting to $1''\cdot75$ outwards.

In either of the last two cases the displacement is inversely proportional to the distance of the star from the sun's centre, the displacement under (3) being just double the displacement under (2).

It may be noted that both (2) and (3) agree in supposing that light is subject to gravitation in precisely the same way as ordinary matter. The difference is that, whereas (2) assumes the Newtonian law, (3) assumes EINSTEIN's new law of gravitation. The slight deviation from the Newtonian law, which on EINSTEIN's theory causes an excess motion of perihelion of Mercury, becomes magnified as the speed increases, until for the limiting velocity of light it doubles the curvature of the path.

2. The displacement (2) was first suggested by Prof. EINSTEIN* in 1911, his argument being based on the Principle of Equivalence, viz., that a gravitational field is indistinguishable from a spurious field of force produced by an acceleration of the axes of reference. But apart from the validity of the general Principle of Equivalence there were reasons for expecting that the electromagnetic energy of a beam of light would be subject to gravitation, especially when it was proved that the energy of radio-activity contained in uranium was subject to gravitation. In 1915, however, EINSTEIN found that the general Principle of Equivalence necessitates a modification of the Newtonian law of gravitation, and that the new law leads to the displacement (3).

3. The only opportunity of observing these possible deflections is afforded by a ray of light from a star passing near the sun. (The maximum deflection by Jupiter is only $0''\cdot017$.) Evidently, the observation must be made during a total eclipse of the sun.

Immediately after EINSTEIN's first suggestion, the matter was taken up by Dr. E. FREUNDLICH, who attempted to collect information from eclipse plates already taken; but he did not secure sufficient material. At ensuing eclipses plans were made by various observers for testing the effect, but they failed through cloud or other causes. After EINSTEIN's second suggestion had appeared, the Lick Observatory expedition attempted to observe the effect at the eclipse of 1918. The final results are not yet published. Some account of a preliminary discussion has been given,† but the eclipse was an unfavourable one, and from the information published the probable accidental error is large, so that the accuracy is insufficient to discriminate between the three alternatives.

4. The results of the observations here described appear to point quite definitely to the third alternative, and confirm EINSTEIN's generalised relativity theory. As is well-known the theory is also confirmed by the motion of the perihelion of Mercury, which exceeds the Newtonian value by 43″ per century—an amount practically identical with that deduced from EINSTEIN's theory. On the other hand, his theory predicts a displacement to the red of the Fraunhofer lines on the sun amounting to about 0·008 Å in the violet. According to Dr. ST. JOHN‡ this displacement is not confirmed. If this disagreement is to be taken as final it necessitates considerable modifications of EINSTEIN's theory, which it is outside our province to discuss. But, whether or not changes are needed in other parts of the theory, it appears now to be established that EINSTEIN's law of gravitation gives the true deviations from the Newtonian law both for the relatively slow-moving planet Mercury and for the fast-moving waves of light.

It seems clear that the effect here found must be attributed to the sun's gravitational field and not, for example, to refraction by coronal matter. In order to produce the observed effect by refraction, the sun must be surrounded by material of refractive index $1 + \cdot00000414/r$, where r is the distance from the centre in terms of the sun's radius. At a height of one radius above the surface the necessary refractive index 1·00000212 corresponds to that of air at $\frac{1}{140}$ atmosphere, hydrogen at $\frac{1}{60}$ atmosphere, or helium at $\frac{1}{20}$ atmospheric pressure. Clearly a density of this order is out of the question.

---

* 'Annalen der Physik,' vol. XXXV, p. 898.

† 'Observatory,' vol. XLII, p. 298.

‡ 'Astrophysical Journal,' vol. XLVI, p. 249.

II. PREPARATIONS FOR THE EXPEDITIONS.

5. In March, 1917,* it was pointed out as the result of an examination of the photographs taken with the Greenwich astrographic telescope at the eclipse of 1905 that this instrument was suitable for the photography of the field of stars surrounding the sun in a total eclipse. Attention was also drawn to the importance of observing the eclipse of May 29, 1919, as this afforded a specially favourable opportunity owing to the unusual number of bright stars in the field, such as would not occur again for many years.

With weather conditions as good as those at Sfax in the 1905 eclipse—and these were by no means perfect—it was anticipated that twelve stars would be shown. Their positions are indicated in the diagram on next page, on which is also marked on the same scale the outline of a 16 × 16 cm. plate (used with the astrographic telescopes of 3·43 metres focal length) and a 10 × 8-inch plate (used with a 4-inch lens of 19 feet focal length).

The following table gives the photographic magnitudes and standard co-ordinates of the stars, and the gravitational displacements in x and y calculated on the assumption of a radial displacement $1''\cdot75\,\frac{r_0}{r}$, where r is the distance from the sun's centre and r_0 the radius of the sun.

TABLE I.

No.	Names.	Photog. Mag.	Co-ordinates. Unit = 50′.		Gravitational displacement.			
					Sobral.		Principe.	
			x.	y.	x.	y.	x.	y.
		m.			″	″	″	″
1	B.D., 21°, 641	7·0	+0·026	−0·200	−1·31	+0·20	−1·04	+0·09
2	Piazzi, IV, 82	5·8	+1·079	−0·328	+0·85	−0·09	+1·02	−0·16
3	κ^2 Tauri	5·5	+0·348	+0·360	−0·12	+0·87	−0·28	+0·81
4	κ^1 Tauri	4·5	+0·334	+0·472	−0·10	+0·73	−0·21	+0·70
5	Piazzi, IV, 61	6·0	−0·160	−1·107	−0·31	−0·43	−0·31	−0·38
6	υ Tauri	4·5	+0·587	+1·099	+0·04	+0·40	+0·01	+0·41
7	B.D., 20°, 741	7·0	−0·707	−0·864	−0·38	−0·20	−0·35	−0·17
8	B.D., 20°, 740	7·0	−0·727	−1·040	−0·33	−0·22	−0·29	−0·20
9	Piazzi, IV, 53	7·0	−0·483	−1·303	−0·26	−0·30	−0·26	−0·27
10	72 Tauri	5·5	+0·860	+1·321	+0·09	+0·32	+0·07	+0·34
11	66 Tauri	5·5	−1·261	−0·160	−0·32	+0·02	−0·30	+0·01
12	53 Tauri	5·5	−1·311	−0·918	−0·28	−0·10	−0·26	−0·09
13	B.D., 22°, 688	8·0	+0·089	+1·007	−0·17	+0·40	−0·14	+0·39

* 'Monthly Notices, R.A.S.,' LXXVII, p. 445.

It may be noted that No. 1 is lost in the corona on the photographs taken at Sobral. The star, No. 13, of magnitude 8·0, is shown on some of the astrographic plates at Sobral.

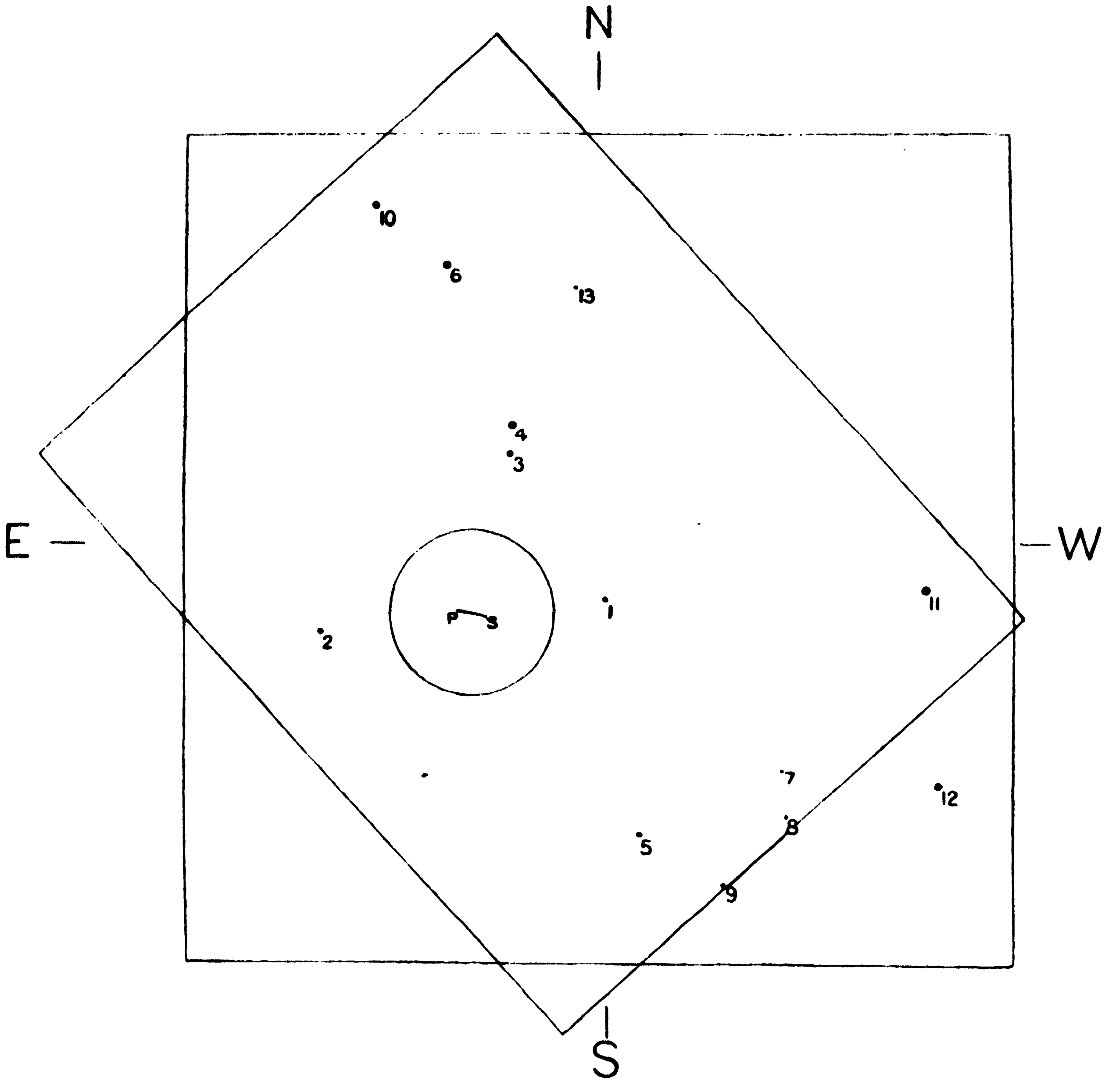


Diagram 1.

6. The track of the eclipse runs from North Brazil across the Atlantic, skirting the African coast near Cape Palmas, passing through the Island of Principe, then across Africa to the western shores of Lake Tanganyika. Enquiry as to the suitable sites and probable weather conditions was kindly made by Mr. Hinks. It appeared that a station in North Brazil, the Island of Principe, and a station on the west of Lake Tanganyika were possible. A station near Cape Palmas did not seem desirable from the meteorological reports though, as the event proved, the eclipse was observed in a cloudless sky

by Prof. BAUER, who was there on an expedition to observe magnetic effects. At the station at Tanganyika it was thought the sun was at too low an altitude for observations of this character, owing to the large displacements which would be caused by refraction.

A circular received from Dr. MORIZE, the director of the Observatory at Rio, stated that Sobral was the most suitable station in North Brazil and gave copious information of the meteorological conditions, mode of access, &c.

7. Acting on this information the Joint Permanent Eclipse Committee at a meeting on November 10, 1917, decided, if possible, to send expeditions to Sobral in North Brazil, and to the island of Principe. Application was made to the Government Grant Committee for £100 for instruments and £1,000 for the expedition, and a sub-committee consisting of Sir F. W. DYSON, Prof. EDDINGTON, Prof. FOWLER and Prof. TURNER was appointed to make arrangements for the expeditions. This sub-committee met in May and June, 1918, and made provisional arrangements for Prof. EDDINGTON and Mr. COTTINGHAM to take the object glass of the Oxford astrographic telescope to Principe, and Mr. DAVIDSON and Father CORTIE to take the object glass of the Greenwich astrographic telescope to Sobral. It was arranged for the clocks and mechanism of the cœlostats to be overhauled by Mr. COTTINGHAM. Preliminary inquiries were also set on foot as to shipping facilities, from which it appeared very doubtful whether the expeditions could be carried through.

Conditions had changed materially in November, 1918, and at a meeting of the sub-committee on November 8, it was arranged to assemble the instruments at Greenwich, and make necessary arrangements with all speed for the observers to leave England by the end of February, 1919. In addition to the astrographic object glasses fed by 16-inch cœlostats, Father CORTIE suggested to the sub-committee the use of the 4-inch telescope of 19-feet focus, which he had used at Hernosand, Sweden, in 1914, in conjunction with an 8-inch cœlostat, the property of the Royal Irish Academy. It was arranged to ask for the loan of these instruments. As Father CORTIE found it impossible to spare the necessary time for the expedition his place was taken by Dr. CROMMELIN of the Royal Observatory.

8. In November, 1918, the only workman available at the Royal Observatory was the mechanic, the carpenter not having been released from military service. In these circumstances Mr. BOWEN, the civil engineer at the Royal Naval College, was consulted. He kindly undertook the construction of frame huts covered with canvas, which could be easily packed and readily put together. These were generally similar to those used in previous expeditions from the Royal Observatory (see 'Monthly Notices,' Vol. LVII., p. 101). He also lent the services of a joiner who worked at the Observatory on the woodwork of the instruments.

It was found possible to obtain steel tubes for the astrographic objectives. These were, for convenience of carriage, made in two sections which could be bolted together. The tubes were provided with flanges at each end, the objective being attached to one of these, and a wooden breech piece to the other. In the breech piece suitable provision was made for the focussing and squaring on of the plates. The plate holders were of a simple construction, permitting the plate to be pushed into contact with three metal tilting screws on the breech piece thus insuring a constancy of focal plane. Eighteen plate-carriers were obtained for each of the astrographic telescopes, made according to a pattern supplied.

With the 4-inch lens Father CORTIE lent the square wooden tube used by him in 1914. This was modified at the breech end to secure greater rigidity and constancy of focus.

It was designed for dark slides carrying 10 × 8 inch plates, and four of these, carrying eight plates, were lent with the telescope. The desirability of using larger plates was considered, but the time at disposal to make the necessary alterations was insufficient.

The 16-inch cœlostats which had been overhauled by Mr. COTTINGHAM were mounted and tested as far as the unfavourable weather conditions of February, 1919, would permit. The 8-inch cœlostat was constructed for these latitudes. To make it serviceable near the equator a strong wooden wedge was made on which the cœlostat was bolted.

The 8-inch mirror was silvered at the observatory, but owing to lack of facilities for maintaining a uniform temperature approaching 60° F. in the wintry weather of February, the larger mirrors were sent away to be silvered.

Photographic plates, suitably packed in hermetically sealed tin boxes, were obtained from the Ilford and Imperial Companies. The Ilford plates employed were Special Rapid and Empress, and those of the Imperial Company, Special Sensitive, Sovereign and Ordinary.

The instruments were carefully packed and sent to Liverpool a week in advance, with the exception of the objectives. These were packed in cases inside hampers and remained under the personal care of the observers, who embarked on the " Anselm " on March 8.

III. THE EXPEDITION TO SOBRAL.

(*Observers, Dr.* A. C. D. CROMMELIN *and Mr.* C. DAVIDSON.)

9. Sobral is the second town of the State of Ceara, in the north of Brazil. Its geographical co-ordinates are: longitude 2h. 47m. 25s. west; latitude 3° 41′ 33″ south; altitude 230 feet. Its climate is dry and though hot not unhealthy.

The expedition reached Para on the " Anselm " on March 23. There was a choice of proceeding immediately to Sobral or waiting for some weeks. It was considered undesirable to go there before we heard from Dr. MORIZE what arrangements were being made, so we reported our arrival to him by telegram and decided to await his reply. As we had thus some time on our hands we continued the voyage to Manaos in the " Anselm," returning to Para on April 8.

By the courtesy of the Brazilian Government our heavy baggage was passed through the customs without examination and we continued our journey to Sobral, leaving Para on April 24 by the steamer " Fortaleza " and arriving at Camocim on April 29.

Here we were met by Mr. JOHN NICOLAU, who had been instructed to assist us with our baggage through to Sobral. We proceeded from Camocim to Sobral by train on April 30, our baggage following the next day.

We were met at Sobral station by representatives of both the Civil and Ecclesiastical Authorities, headed respectively by Dr. JACOME D'OLIVEIRA, the Prefect, and Mgr. FERREIRA, and conducted to the house which had been placed at our disposal by the kindness of its owner, Col. VICENTE SABOYA, the Deputy for Sobral. We were joined there nine days later by the Washington (Carnegie) Eclipse Commission, consisting of Messrs. DANIEL WISE and ANDREW THOMSON.

We are greatly indebted to Dr. LEOCADIO ARAUJO, of the State Ministry of Agriculture, who had been deputed to interpret for us and to assist us in our preparations. His services were invaluable, and contributed greatly to our success, as also to our well-being during our stay.

10. A convenient site for the eclipse station offered itself just in front of the house; this was the race-course of the Jockey Club, and was provided with a covered grand stand, which we found most convenient for unpacking and storage and in the preparatory work. We laid down a meridian line, after which brick piers were constructed for the cœlostats and for the steel tube of the astrographic telescope. Whilst this was in progress the huts were being erected.

The pier of the small cœlostat was constructed so as to leave a clear space in the middle of one end for the fall of the weight, which was thus below the driving barrel of the clock. By continuing the hole below the foundations of the pier, space was provided for a fall of the weight permitting a run of 25 minutes. In the case of the 16-inch cœlostat, the clock was mounted on the top of a long wooden trunk, nearly 4 feet in length, which was placed on end, and sunk in the earth to a depth of about 2 feet. The weight descended inside the trunk directly from the driving barrel, and had space for a continuous run of over half-an-hour.

The 16-inch cœlostat had free adjustment for all latitudes; but the 8-inch one, constructed for European latitudes, was mounted on a wooden base, inclined at an angle of about 40 degrees, constructed before leaving Greenwich. The clock had to be separated from the cœlostat, mounted on a wooden base and reversed, to adjust to the Southern Hemisphere. It performed very satisfactorily, and no elongation of the star images is shown with 28 seconds' exposure.

To provide for the changing declination of the sun the piers of the astrographic telescope were made with grooves in the top, in which the wooden V-supports of the tube could slide, thus allowing for the change of azimuth.

The tube of the astrographic telescope was circular in section, and could rest in any position in the Vs; for convenience it was adjusted so that the directions of R.A. and declination were parallel to the sides of the plate; this involved a tilt of the plate holders of about 4 degrees to the horizontal.

The 4-inch lens was taken as an auxiliary; we used the square wooden tube, 19 feet in length, originally used by Father CORTIE at Hernosand in 1914, together with the 10 × 8-inch plate carriers. Study of the star-diagram showed that seven stars could be photographed by turning the plate through 45 degrees. The tube was therefore placed on its angle, large wooden V-supports being prepared to fit the tube; these rested on strong wooden trestles.

The focussing was at first done visually on Arcturus, using an eyepiece fitted with a cobalt glass (after the plate supports and object-glass had been adjusted for perpendicularity to the axis). A series of exposures was then made, the focus being varied slightly so as to cover a sufficient range. Examination of these photographs showed at once that there was serious astigmatism due to the figure of the mirror of the 16-inch cœlostat. By inserting an 8-inch stop this was reduced to a large extent, and this stop was henceforth used throughout; but the defect was of such a character that it was clear that it would be necessary to stay at Sobral and obtain comparison plates of the eclipse field in July when the sun had moved away.

The focus of the 4-inch was determined in a similar manner. The images, though superior to those of the astrographic, were not quite perfect, and here again comparison plates in July were necessary. Once the focus had been decided on, the breech end was securely screwed up to avoid any chance of subsequent movement.

A few check plates of the field near Arcturus were taken, but have not been used.

11. The following is a summary of the meteorological conditions during our stay. The barometer record was interesting in that it showed very little change from day to day, in spite of changes in the type of weather; there was, however, a very well marked semi-diurnal variation, with range of about 0·15 inch. The temperature range was fairly uniform, from a maximum of about 97° F. towards 3 p.m. to a minimum of about 75° F. at 5 a.m. The relative humidity (as shown by a hygrograph belonging to the Brazilian Commission) followed the temperature closely, varying from 30 per cent. in the afternoon to 90 per cent. in the early morning.

May is normally the last month of the rainy season at Sobral, but this year the rainfall was very scanty; there were a few afternoon showers, each ushered in by a violent gust of wind; and on May 25 there was very heavy rain, which was welcome for its moistening effect on the ground, the dust hitherto having been troublesome to the clockwork although every care had been taken to protect it. There was a fair amount of cloud in the mornings, but the afternoons and nights were clear in the majority of cases. Mt. Meruoca, 2,700 feet high, about 6 miles to the N.W., was a collector of cloud, its summit being frequently veiled in mist. In spite of its cooler climate, the summit would thus not have been a suitable eclipse station, and, in fact, nothing of the total phase of the eclipse was seen from it.

12. Although water was generally scarce, we were very fortunately situated as we enjoyed an unlimited supply of good water laid on at the house. This was of great benefit in the photographic operations. Ice was unobtainable, but by the use of earthenware water-coolers it was possible to reduce the temperature to about 75°, and by working

only at night or before dawn development of the plates was fairly easy. Formalin was used in every case to harden the films, and thereby minimise the chance of distortion due to the softening of the films by the warm solutions.

We had provided ourselves with two brands of plates, but it had become apparent from photographs taken and developed before the eclipse that one of these brands was unsuitable in the hot climate, and it was decided to use practically only one brand of plates.

In taking the experimental photographs it was noticed that the clocks and cœlostats were very sensitive to wind. We had reason to fear strong gusts about the time of totality, such as had occurred in other eclipses ; and as the conditions of our locality seemed to render them specially probable, protective wind screens were erected round the hut openings at every point where it was possible without interfering with the field of view. Happily dead calm prevailed at the critical time. Screens also protected all projecting parts of the telescope tubes from direct sunlight.

The performance of the 16-inch cœlostat was unsatisfactory in respect of driving. There was a clearly marked oscillation of the images on the screen in a period of about 30 seconds. For this reason exposure time was shortened, so as to multiply the number of exposures in the hope that some would be near the stationary points.

13. The morning of the eclipse day was rather more cloudy than the average, and the proportion of cloud was estimated at $\frac{9}{10}$ at the time of first contact, when the sun was invisible ; it appeared a few seconds later showing a very small encroachment of the moon, and there were various short intervals of sunshine during the partial phase which enabled us to place the sun's image at its assigned position on the ground glass, and to give a final adjustment to the rates of the driving clocks. As totality approached, the proportion of cloud diminished, and a large clear space reached the sun about one minute before second contact. Warnings were given 58s., 22s. and 12s. before second contact by observing the length of the disappearing crescent on the ground glass. When the crescent disappeared the word "go" was called and a metronome was started by Dr. Leocadio, who called out every tenth beat during totality, and the exposure times were recorded in terms of these beats. It beat 320 times in 310 seconds ; allowance has been made for this rate in the recorded times. The programme arranged was carried out successfully, 19 plates being exposed in the astrographic telescope with alternate exposures of 5 and 10 seconds, and eight in the 4-inch camera with a uniform exposure of 28 seconds. The region round the sun was free from cloud, except for an interval of about a minute near the middle of totality when it was veiled by thin cloud, which prevented the photography of stars, though the inner corona remained visible to the eye and the plates exposed at this time show it and the large prominence excellently defined. The plates remained in their holders until development, which was carried out in convenient batches during the night hours of the following days, being completed by June 5.

14. No observation of contact times was made, but it is known that these times were

somewhat before those calculated. As the times recorded were reckoned from second contact, it is assumed that this occurred May 28, 23h. 58m. 18s. G.M.T.

The details of the exposures are given in the following tables :—

EXPOSURES with the 13-inch Astrographic Telescope stopped to 8 inches.

Ref. No.	G.M.T. at Commencement of Exposure.	Exposure.	Plate.	Ref. No.	G.M.T. at Commencement of Exposure.	Exposure.	Plate.
	d. h. m. s.	s.			d. h. m. s.	s.	
1	28 23 58 23	5	O.	11	29 0 1 7	5	S.R.
2	37	10	E.	12	22	10	E.
3	57	5	E.	13	36	5	E.
4	59 11	10	S.	14	51	10	S.R.
5	30	5	S.R.	15	2 10	5	S.R.
6	45	10	S.R.	16	25	10	S.R.
7	29 0 0 4	5	S.R.	17	44	5	E.
8	19	10	E.	18	58	10	E.
9	39	5	E.	19	3 18	5	O.
10	53	10	S.R.				

EXPOSURES with the 4-inch Telescope.

Ref. No.	G.M.T. at Commencement of Exposure.	Exposure.	Plate.	Ref. No.	G.M.T. at Commencement of Exposure.	Exposure.	Plate.
	d. h. m. s.	s.			d. h. m. s.	s.	
1	28 23 58 21	28	S.R.	5	29 0 0 56	28	S.R.
2	59 0	28	S.R.	6	1 34	28	S.R.
3	38	28	S.R.	7	2 13	28	S.R.
4	29 0 0 17	28	S.R.	8	52	28	S.R.

In the fourth column the letter O stands for Imperial Ordinary.
E ,, ,, Empress.
S ,, ,, Sovereign.
SR ,, ,, Ilford Special Rapid.

With the astrographic telescope 12 stars are shown on a number of plates, and seven stars on all but three (Nos. 13, 14 and 19). Of the eight plates taken with the 4-inch lens, seven show seven stars, but No. 6, which was taken through cloud, does not show any.

The following table of temperatures, communicated by Dr. MORIZE, and converted into the Fahrenheit scale, shows how slight the fall was during totality, probably owing to the large amount of cloud in the earlier stages which checked the usual daily rise.

G.M.T.	Ther.	G.M.T.	Ther.	G.M.T.	Ther.	G.M.T.	Ther.
d. h. m.	°	d. h. m.	°	d. h. m.	°	d. h. m.	°
28 22 45	82·4	28 23 30	80·6	29 0 15	82·0	29 1 0	83·8
23 0	84·2	45	82·4	30	82·4	15	84·2
15	82·4	29 0 0	80·6	45	83·1	30	84·2

15. On June 7, having completed the development, we left Sobral for Fortaleza, returning on July 9 for the purpose of securing comparison plates of the eclipse field.

Before our departure we dismounted the mirrors and driving clocks which were brought into the house to avoid the exposure to dust. The telescopes and cœlostats were left *in situ.* Before removing the mirrors we marked their positions in their cells so that they could be replaced in exactly the same position.

After our return to Sobral the mirrors and clocks were remounted ; the photography of the eclipse field was commenced on the morning of July 11 (civil). The difficulty of finding the field with the cœlostats was overcome by making a rough hour circle on the heads of the cœlostats out of millimetre paper.

The following is the list of exposures made on the field for comparison with the eclipse photographs :—

Astrographic Telescope.						4-inch Telescope.					
Ref. No.	Date.	G.M.T.	No. of exposures.	Duration.	Altitude.	Ref. No.	Date.	G.M.T.	No. of exposures.	Duration.	Altitude.
		h. m.		s.	°						
11_1	July 10	20 5	3	5	28·9						
11_2		20 16	2	5	31·1			h. m.		s.	°
11_3		20 21	1	5	32·2	14_1	July 13	20 7	2	25	32·4
14_1	July 13	20 13	3	5	33·7	14_2		20 16	2	20	34·3
14_2		20 17	2	5	34·5						
14_3		20 19	2	5	34·9	15_1	July 14	20 17	2	20	35·4
15_1	July 14	20 15	3	5	34·9	15_2		20 22	2	20	36·4
15_2		20 20	2	5	36·1						
15_3		20 23	2	5	36·6	17_1	July 16	20 6	3	15	34·7
17_1	July 16	20 2	4	3	33·8	17_2		20 24	2	15	38·6
17_2		20 15	3	3	36·6						
17_3		20 23	2	3	38·3						
17_4		20 25	2	5	38·8	18_1*	July 17	19 57	3	20	33·6
18_1	July 17	19 50	3	4	32·8	18_2		20 24	2	20	39·2
18_2		20 1	2	4	34·4						
18_3		20 20	3	4	38·6						
18_4		20 25	2	3	39·5						

The reference numbers follow the civil dates.

* The 4-inch plate, No. 18_1, was taken through the glass (see § 17, *infra*) to facilitate the measurement, and is referred to as the scale plate.

Thermometer readings, July 10, 74°·4; July 13, 73°·7; July 14, 71°·9; July 16, 72°·3; July 17, 72°·3.

By July 18 we had obtained a sufficient number of reference photographs. Dismantling of the instruments was commenced, and the packing was completed on July 21. We left Sobral on July 22, leaving the packing cases in the hands of Messrs. NICOLAU and CARNEIRO to be forwarded at the earliest opportunity, and arrived at Greenwich on August 25.

The observers wish to record their obligations to Mr. CHARLES BOOTH and the officers of the "Booth" Line for facilitating their journeys to and from their station at a difficult time.

PHOTOGRAPHS TAKEN WITH THE 4-INCH OBJECT GLASS.

16. These photographs were taken on 10 × 8-inch plates. By suitably mounting the camera it was made possible to obtain seven stars on the photographs, viz., Nos. 2, 3, 4, 5, 6, 10 and 11 of the table in § 5. Of the eight photographs taken during the eclipse seven gave measurable images of these stars, the other plate (No. 6) taken through cloud only showing a picture of the prominences.

Plates of the same field taken under nearly similar conditions as regards altitude were taken on July 14, 15, 17 and 18 (civil date). Of these photographs, the second taken on July 14 with two exposures (referred to as 14_{2a} and 14_{2b}), two photographs taken on July 15 (referred to as 15_1 and 15_2), two on July 17 (17_1 and 17_2), and the second photograph on July 18 (18_2) were measured for comparison with the eclipse plates.

17. The micrometer at the Royal Observatory is not suitable for the direct comparison of plates of this size. It was therefore decided to measure each plate by placing, film to film upon it, another photograph of the same region reversed by being taken through the glass. A photograph for this purpose was taken on July 18. This plate is regarded merely as an intermediary between the eclipse plates and comparison plates and is referred to as the scale plate, being used simply as a scale providing points of reference. In all cases measurement was made through the glass of the scale plate, adjusted on the eclipse or comparison plate which was being measured, so that the separation of the images on the two plates did not exceed one-third of a millimetre. The plates were held together by clips which ensured contact over the whole surface. This method of measurement was found to be very convenient. Each plate was measured in two positions, being reversed through 180 degrees, and the accordance of the result showed that the method of measurement was entirely satisfactory.

The measures, both direct and reversed, were made by two measurers (Mr. DAVIDSON and Mr. FURNER), and the means taken. There was no sensible difference between the measurers, which is satisfactory, as it affords evidence of the similarity of the images on the eclipse and comparison and scale plates.

The value of the micrometer screws (both in R.A. and Decl.) is 6″·25.

18. The results of the measures are as follows :—

TABLE II.—Eclipse Plates—Scale.

No. of Star.	I.		II.		III.		IV.		V.		VII.		VIII.	
	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.
	r	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*
11	—1·411	—0·554	—1·416	—1·324	+0·592	+0·956	+0·563	+1·238	+0·406	+0·970	—1·456	+0·964	—1·285	—1·195
5	—1·048	—0·338	—1·221	—1·312	+0·756	+0·843	+0·683	+1·226	+0·468	+0·861	—1·267	+0·777	—1·152	—1·332
4	—1·216	+0·114	—1·054	—0·944	+0·979	+1·172	+0·849	+1·524	+0·721	+1·167	—1·028	+1·142	—0·927	—0·930
3	—1·237	+0·150	—1·079	—0·862	+0·958	+1·244	+0·861	+1·587	+0·733	+1·234	—1·010	+1·185	—0·897	—0·894
6	—1·342	+0·124	—1·012	—0·932	+1·052	+1·197	+0·894	+1·564	+0·798	+1·130	—0·888	+1·125	—0·838	—0·937
10	—1·289	+0·205	—0·999	—0·948	+1·157	+1·211	+0·934	+1·522	+0·864	+1·119	—0·820	+1·072	—0·768	—0·964
2	—0·789	+0·109	—0·733	—1·019	+1·256	+0·924	+1·177	+1·373	+0·995	+0·935	—0·768	+0·892	—0·585	—1·166
	—1·500*	—0·554	—1·500	—1·324	+0·500	+0·843	+0·500	+1·226	+0·400	+0·861	—1·500	+0·777	—1·300	—1·322

COMPARISON Plates—Scale.

No. of Star.	14_{2a}.		14_{2b}.		15_1.		15_2.		17_1.		17_2.		18_2.	
	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.	D*x*.	D*y*.
	r	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*
11	—0·478	—0·109	+0·967	+1·170	+1·098	+1·228	+0·725	+0·830	—1·073	—1·330	+1·242	—0·302	—1·188	—1·572
5	—0·544	—0·204	+1·013	+1·192	+0·899	+1·232	+0·692	+0·938	—1·072	—1·075	+1·161	—0·224	—1·195	—1·432
4	—0·368	—0·136	+1·030	+1·249	+1·133	+1·086	+0·725	+0·854	—1·296	—1·031	+1·354	—0·281	—1·165	—1·454
3	—0·350	—0·073	+1·044	+1·305	+1·164	+1·114	+0·732	+0·893	—1·278	—1·014	+1·342	—0·261	—1·178	—1·394
6	—0·317	—0·144	+0·980	+1·319	+1·244	+1·012	+0·714	+0·824	—1·375	—1·052	+1·363	—0·390	—1·165	—1·473
10	—0·272	—0·146	+0·997	+1·327	+1·249	+0·960	+0·722	+0·831	—1·424	—1·038	+1·370	—0·423	—1·164	—1·476
2	—0·396	—0·182	+1·102	+1·289	+0·969	+1·052	+0·734	+0·941	—1·236	—0·909	+1·278	—0·328	—1·164	—1·335
	—0·552*	—0·206	+0·967	+1·170	+0·899	+0·960	+0·690	+0·824	—1·424	—1·330	+1·161	—0·423	—1·195	—1·572

---

* The numbers —1·500, —0·554, &c., given below the line, were taken out to make the values of D*x*, D*y* small and positive for arithmetical convenience.

19. The values of Dx and Dy were equated to expressions of the form

$$ax + by + c + \alpha E_x \ (= Dx)$$

and

$$dx + ey + f + \alpha E_y \ (= Dy),$$

where x, y are the co-ordinates of the stars given in Table I., and E_x, E_y are coefficients of the gravitational displacement.

The quantities c and f are corrections to zero, depending on the setting of the scale plate on the plate measured, a and e are differences of scale value, while b and d depend mainly on the orientation of the two plates. The quantity α denotes the deflection at unit distance (*i.e.*, 50′ from the sun's centre), so that αE_x and αE_y are the deflection in R.A. and Decl. respectively of a star whose co-ordinates are x and y.

The left-hand sides of the equation for the seven stars shown are :—

No.	Right Ascension.	Declination.
11	$c-0\cdot160b-1\cdot261a-0\cdot587\alpha$	$f-1\cdot261d-0\cdot160e+0\cdot036\alpha$
5	$c-1\cdot107b-0\cdot160a-0\cdot557\alpha$	$f-0\cdot160d-1\cdot107e-0\cdot789\alpha$
4	$c+0\cdot472b+0\cdot334a-0\cdot186\alpha$	$f+0\cdot334d+0\cdot472e+1\cdot336\alpha$
3	$c+0\cdot360b+0\cdot318a-0\cdot222\alpha$	$f+0\cdot348d+0\cdot360e+1\cdot574\alpha$
6	$c+1\cdot099b+0\cdot587a+0\cdot080\alpha$	$f+0\cdot587d+1\cdot099e+0\cdot726\alpha$
10	$c+1\cdot321b+0\cdot860a+0\cdot158\alpha$	$f+0\cdot860d+1\cdot321e+0\cdot589\alpha$
2	$c-0\cdot328b+1\cdot079a+1\cdot540\alpha$	$f+1\cdot079d-0\cdot328e-0\cdot156\alpha$

20. Normal equations formed from these equations of condition are as follows :—

TABLE III.—Eclipse Plates—Right Ascension.

					I.	II.	III.	IV.	V.	VII.	VIII.
$+7\cdot000c$	$+1\cdot657b$	$+1\cdot787a$	$+0\cdot226\alpha$	=	+2·159	+2·986	+3·250	+2·461	+2·185	+3·263	+2·648
	$+4\cdot664$	$+2\cdot089$	$+0\cdot335$	=	−0·063	+0·986	+1·320	+0·866	+1·051	+1·464	+1·130
		$+4\cdot094$	$+2\cdot534$	=	+1·034	+1·689	+1·866	+1·469	+1·480	+1·972	+1·723
			$+3\cdot142$	=	+0·712	+0·919	+0·924	+0·860	+0·844	+0·930	+0·973
	$+4\cdot271b$	$+1\cdot666a$	$+0\cdot281\alpha$	=	−0·575	+0·278	+0·550	+0·283	+0·533	+0·691	+0·502
		$+3\cdot683$	$+2\cdot476$	=	+0·483	+0·928	+1·037	+0·841	+0·923	+1·140	+1·048
			$+3\cdot135$	=	+0·643	+0·823	+0·820	+0·781	+0·774	+0·826	+0·888
		$+2\cdot988a$	$+2\cdot366\alpha$	=	+0·707	+0·820	+0·822	+0·731	+0·715	+0·871	+0·852
			$+3\cdot116$	=	+0·681	+0·805	+0·784	+0·762	+0·739	+0·780	+0·855
			$+1\cdot242\alpha$	=	+0·121	+0·156	+0·133	+0·183	+0·173	+0·090	+0·180
			α	=	+0·098	+0·126	+ ·107	+0·148	+0·140	+0·073	+0·145
			a	=	+0·158	+0·174	+0·189	+0·127	+0·128	+0·233	+0·169
			b	=	−0·203	−0·011	+0·048	+0·007	+0·042	+0·066	+0·042

TABLE IV.—Comparison Plates—Right Ascension.

					14_{2a}.	14_{2b}.	15_1.	15_2.	17_1.	17_2.	18_2.
+7·000*c*	+1·657*b*	+1·787*a*	+0·226α	=	+1·190	+0·364	+1·463	+0·214	+1·214	+0·983	+0·146
	+4·664	+2·089	+0·335	=	+0·700	+0·017	+0·992	+0·078	−0·340	+0·603	+0·083
		+4·094	+2·535	=	+0·638	+0·220	+0·499	+0·073	−0·172	+0·450	+0·085
			+3·142	=	+0·253	+0·159	−0·029	+0·037	−0·164	+0·105	+0·041
	+4·271*b*	+1·666*a*	+0·281α	=	+0·418	−0·069	+0·645	+0·027	−0·627	+0·370	+0·048
		+3·683	+2·476	=	+0·334	+0·127	+0·126	+0·018	−0·481	+0·199	+0·048
			+3·135	=	+0·215	+0·147	−0·076	+0·030	−0·203	+0·074	+0·036
		+2·988*a*	+2·366α	=	+0·172	+0·154	−0·126	+0·007	−0·236	+0·055	+0·029
			+3·116	=	+0·188	+0·152	−0·119	+0·028	−0·162	+0·050	+0·033
			+1·242α	=	+0·052	+0·030	−0·019	+0·022	+0·025	+0·006	+0·010
			α	=	+0·042	+0·024	−0·015	+0·018	+0·020	+0·005	+0·008
			a	=	+0·024	+0·032	−0·030	−0·012	−0·094	+0·014	+0·003
			b	=	+0·086	−0·030	+0·164	+0·012	−0·111	+0·081	+0·010

TABLE V.—Eclipse Plates—Declination.

+7·000*f*	+1·787*d*	+1·657*e*	+3·316α	=	+3·688	+1·927	+1·646	+1·452	+1·389	+1·718	+1·906
	+4·094	+2·089	+1·840	=	+2·200	+1·168	+0·719	+0·823	+0·555	+0·610	+0·840
		+4·664	+3·694	=	+1·860	+1·159	+1·129	+0·984	+0·874	+1·023	+1·193
			+5·784	=	+2·657	+1·681	+1·535	+1·361	+1·335	+1·545	+1·707
	+3·638*d*	+1·666*e*	+0·994α	=	+1·260	+0·677	+0·299	+0·453	+0·201	+0·172	+0·354
		+4·271	+2·908	=	+0·986	+0·702	+0·739	+0·640	+0·545	+0·616	+0·741
			+4·212	=	+0·909	+0·768	+0·755	+0·673	+0·677	+0·731	+0·804
		+3·508*e*	+2·453α	=	+0·409	+0·392	+0·602	+0·431	+0·453	+0·537	+0·579
			+3·941	=	+0·565	+0·583	+0·673	+0·549	+0·622	+0·684	+0·707
			+2·224α	=	+0·279	+0·309	+0·252	+0·247	+0·305	+0·308	+0·302
			α	=	+0·126	+0·139	+0·114	+0·111	+0·137	+0·139	+0·136
			e	=	+0·029	+0·015	+0·092	+0·045	+0·033	+0·056	+0·070
			d	=	+0·299	+0·141	+0·009	+0·074	+0·003	−0·016	+0·028

TABLE VI.—Comparison Plates—Declination.

+7·000*f*	+1·787*d*	+1·657*e*	+3·316α	=	+0·446	+0·661	+0·964	+0·343	+1·861	+0·752	+0·868
	+4·094	+2·089	+1·840	=	+0·060	+0·420	−0·156	+0·140	+1·038	+0·041	+0·476
		+4·664	+3·694	=	+0·202	+0·394	−0·203	−0·117	+0·526	−0·110	+0·122
			+5·784	=	+0·380	+0·482	+0·220	+0·044	+1·004	+0·296	+0·419
	+3·638*d*	+1·666*e*	+0·994α	=	−0·054	+0·251	−0·402	+0·053	+0·563	+0·151	+0·255
		+4·271	+2·908	=	+0·096	+0·237	−0·431	−0·198	+0·085	−0·288	−0·084
			+4·212	=	+0·168	+0·169	−0·237	−0·119	+0·122	−0·060	+0·008
		+3·508*e*	+2·453α	=	+0·121	+0·122	−0·247	−0·222	−0·173	−0·219	−0·201
			+3·941	=	+0·183	+0·100	−0·127	−0·133	−0·032	−0·019	−0·062
			+2·224α	=	+0·098	+0·015	+0·046	+0·022	+0·089	+0·134	+0·079
			α	=	+0·044	+0·007	+0·021	+0·010	+0·040	+0·060	+0·036
			e	=	+0·004	+0·030	−0·085	−0·070	−0·077	−0·104	−0·082
			d	=	−0·028	+0·054	−0·077	+0·044	+0·179	−0·010	+0·098

21. The values of α are collected in Table VII :—

TABLE VII.

Right Ascension.		Declination.	
Eclipse — Scale.	Comparison — Scale.	Eclipse — Scale.	Comparison — Scale.
r	r	r	r
+0·098	+0·042	+0·126	+0·044
+0·126	+0·024	+0·139	+0·007
+0·107	−0·015	+0·114	+0·021
+0·148	+0·018	+0·111	+0·010
+0·140	+0·020	+0·137	+0·040
+0·073	+0·005	+0·139	+0·060
+0·145	+0·008	+0·136	+0·036
Mean +0·120	+0·015	+0·129	+0·031

By subtracting the α of the comparison plates the scale plate is eliminated, and we derive from right ascensions $\alpha = +0^r{\cdot}105$ and from declinations $\alpha = +0^r{\cdot}098$.

Reference to the normal equations shows that the declination result is of double the weight of that from the right ascensions.

Thus

$$\alpha = +0^r{\cdot}100 = +0''{\cdot}625.$$

This is at a distance 50′ from the sun's centre. At the time of the eclipse the sun's radius was 15′·8 ; thus the deflection at the limb is 1″·98.

The range in the values of α is attributable to the errors inherent to the star images of the different plates, and cannot be reduced by further measurement. The mean values $+0^r{\cdot}015$ and $0^r{\cdot}031$ arise from the errors in the intermediary scale plate.

22. The probable error of the result judging from the accordance of the separate determinations is about 6 per cent. It is desirable to consider carefully the possibility of systematic error. The eclipse and comparison photographs were taken under precisely similar instrumental conditions, but there is the difference that the eclipse photographs were taken on the day of May 29, and the comparison photographs on nights between July 14 and July 18. A very satisfactory feature of the photographs is the essential similarity of the star images on the two sets of photographs.

The satisfactory accordance of the eclipse and comparison plates is shown by a study of the plate constants. The following corrections for differential refraction and aberration are calculated from the times and dates of exposure.

	a.	*e.*	*b.*	*d.*
	r	*r*	*r*	*r*
Eclipse plates	+0·240	+0·168	+0·062	+0·062
Scale plate	+0·423	+0·207	+0·096	+0·096
Comparison 14_{2a}	+0·409	+0·207	+0·091	+0·091
,, 14_{2b}	+0·409	+0·207	+0·091	+0·091
,, 15_1	+0·390	+0·207	+0·087	+0·087
,, 15_2	+0·370	+0·202	+0·087	+0·087
,, 17_1	+0·399	+0·207	+0·091	+0·091
,, 17_2	+0·337	+0·202	+0·077	+0·077
,, 18_2	+0·327	+0·202	+0·072	+0·072

When these are applied to the values of the constants found from the normal equations, we find the following values of the scale of the several photographs and their orientation relative to the scale plate :—

	Scale Value.		Orientation.		Adopted Scale Orientation.	
	From *x*.	From *y*.	From *x*.	From *y*.		
	r	*r*				*r*
Eclipse I.	−0·025	−0·010	−0·237	−0·265	0·000	−0·251
,, II.	−0·009	−0·024	−0·045	−0·107	0·000	−0·076
,, III.	+0·006	+0·053	+0·014	+0·025	0·000	+0·020
,, IV.	−0·056	+0·006	−0·027	−0·040	0·000	−0·034
,, V.	−0·055	−0·006	+0·008	+0·031	0·000	+0·020
,, VII.	+0·050	+0·017	+0·032	+0·050	0·000	+0·041
,, VIII.	−0·014	+0·031	+0·008	+0·006	0·000	+0·007
Comparison 14_{2a}	+0·010	+0·004	+0·081	+0·033	+0·013	+0·057
,, 14_{2b}	+0·008	+0·030	−0·035	−0·049	+0·013	−0·042
,, 15_1	−0·063	−0·085	+0·155	+0·086	−0·084	+0·120
,, 15_2	−0·065	−0·075	+0·003	−0·035	−0·084	−0·016
,, 17_1	−0·118	−0·077	−0·116	−0·174	−0·084	−0·145
,, 17_2	−0·072	−0·109	+0·062	+0·029	−0·084	+0·046
,, 18_2	−0·093	−0·087	−0·014	−0·074	−0·084	−0·044

The agreement in the scale values obtained from *x* and *y* is satisfactory. There appears to be a small difference in the orientations as derived from the two directions in the comparison plates. This is, however, of small importance in the determination of α. There is a difference of scale value from July 15–18 shown in both co-ordinates. For the purpose of exhibiting the gravitational displacements, residuals have been computed using adopted values for the scale and orientation given above, along with the calculated corrections for differential refraction and aberration. This has the advantage of reducing the number of constants employed in the reduction of the plates, and lessens the possibility of masking any discordances, though greater irregularities necessarily appear when four arbitrary constants instead of six are used in the reduction of each plate. The quantities are converted from revolutions to seconds of arc, as the more familiar unit facilitates judgment of the results.

TABLE VIII.—Comparison of the Eclipse and Comparison Photographs with the Scale Plate, after Correction for Differential Refraction and Aberration, Orientation and Change of Scale.

No. of Star.	I.	II.	III.	IV.	V.	VII.	VIII.	Mean.
			ECLIPSE Plates—Right Ascension.					
	″	″	″	″	″	″	″	″
11	−0·18	−0·51	−0·46	−0·07	−0·04	−0·72	−0·43	−0·34
5	−0·45	−0·81	−0·38	−0·58	−0·60	−0·36	−0·62	−0·54
4	+0·08	+0·11	−0·08	−0·11	−0·11	−0·16	−0·18	−0·06
3	−0·23	−0·11	−0·19	−0·05	−0·02	−0·02	−0·01	−0·09
6	−0·14	+0·23	−0·09	−0·11	−0·13	+0·13	−0·08	−0·03
10	+0·17	+0·06	+0·14	−0·18	−0·11	+0·14	−0·01	+0·03
2	+0·75	+1·03	+1·06	+1·09	+1·01	+0·98	+1·30	+1·03
			ECLIPSE PLATES—Declination.					
	″	″	″	″	″	″	″	″
11	0·00	−0·08	−0·03	+0·02	+0·17	+0·16	+0·01	+0·03
5	−0·38	−0·54	−0·61	−0·30	−0·39	−0·73	−0·81	−0·54
4	+1·19	+1·04	+1·03	+0·98	+1·11	+1·19	+1·24	+1·11
3	+1·42	+1·58	+1·50	+1·39	+1·55	+1·49	+1·49	+1·49
6	+0·65	+0·79	+1·01	+0·97	+0·71	+0·95	+1·01	+0·87
10	+0·62	+0·46	+1·03	+0·54	+0·56	+0·58	+0·74	+0·65
2	+0·01	+0·25	−0·40	−0·09	−0·22	−0·14	−0·17	−0·11

	14_{2a}.	14_{2b}.	15_1.	15_2.	17_1.	17_2.	18_2.	Mean.
			COMPARISON Plates—Right Ascension.					
	″	″	″	″	″	″	″	″
11	−0·19	−0·24	−0·23	−0·28	+0·11	−0·19	−0·02	−0·15
5	−0·42	+0·16	−0·36	−0·32	−0·24	−0·33	−0·26	−0·25
4	−0·01	+0·03	−0·01	+0·05	−0·04	+0·23	+0·08	+0·05
3	+0·14	+0·09	+0·28	+0·10	−0·03	+0·21	−0·01	+0·11
6	+0·02	−0·18	+0·26	+0·06	+0·13	+0·03	+0·14	+0·07
10	+0·17	−0·06	+0·20	+0·18	+0·13	−0·02	+0·15	+0·11
2	+0·31	+0·18	−0·16	+0·22	−0·04	+0·08	−0·06	+0·08
			COMPARISON Plates—Declination.					
	″	″	″	″	″	″	″	″
11	−0·07	+0·08	−0·26	−0·04	−0·26	−0·18	−0·16	−0·13
5	−0·23	−0·03	+0·03	0·00	−0·19	+0·03	−0·20	−0·08
4	+0·23	+0·05	+0·29	+0·18	+0·45	+0·53	+0·23	+0·28
3	+0·64	+0·41	+0·42	+0·36	+0·48	+0·60	+0·54	+0·49
6	+0·22	+0·36	+0·33	+0·26	+0·41	+0·21	+0·32	+0·30
10	+0·28	+0·32	+0·31	+0·36	+0·36	+0·15	+0·29	+0·30
2	+0·25	+0·14	+0·18	+0·21	+0·09	−0·03	+0·27	+0·16

Subtracting the results of the comparison plates, so as to eliminate the errors arising from the intermediary scale plate we find for the displacements of the different stars, as compared with those as given by EINSTEIN's Theory, with value 1″·75 at the sun's limb :—

No. of Star.	Displacement in Right Ascension.		Displacement in Declination.	
	Observed.	Calculated	Observed.	Calculated.
	″	″	″	″
11	−0·19	−0·32	+0·16	+0·02
5	−0·29	−0·31	−0·46	−0·43
4	−0·11	−0·10	+0·83	+0·74
3	−0·20	−0·12	+1·00	+0·87
6	+0·10	+0·04	+0·57	+0·40
10	−0·08	+0·09	+0·35	+0·32
2	+0·95	+0·85	−0·27	−0·09

PHOTOGRAPHS TAKEN WITH THE ASTROGRAPHIC OBJECT GLASS.

23. As stated above these photographs were taken with the astrographic object glass stopped down to 8 inches, mounted in a steel tube and fed by a 16-inch cœlostat. From many years' experience with the object glass at Greenwich it is certain that, when the object glass is mounted in a steel tube, the change of scale over a range of temperature of 10° F. should be insignificant, and the definition should be very good. It was realised that this high standard would not be obtained with the glass used in conjunction with the cœlostat taken to Brazil, but nevertheless the results shown when the plates were developed were very disappointing. The images were diffused and apparently out of focus, although on the night of May 27 the focus was good.* Worse still, this change was temporary, for without any change in the adjustments, the instrument had returned to focus when the comparison plates were taken in July.

These changes must be attributed to the effect of the sun's heat on the mirror, but it is difficult to say whether this caused a real change of scale in the resulting photographs or merely blurred the images.

The photographs were measured in the astrographic duplex micrometer, the eclipse photographs being directly compared with the comparison plates taken in July. All the stars shown were measured. They were reduced by the same method as that employed for the "4-inch" photographs. With the exception of plates Nos. 15 and 16, taken through clouds, the stars numbered 3, 4, 5, 6, 10, 11 and 12 are shown on all the plates; the fainter stars 2, 7, 8 and 9 are sometimes shown, but No. 1, which is very near the sun, is always drowned in the corona. These plates were only measured in declination, as the right ascensions were of little weight.

---

* The following note made at the time is quoted in full :—" May 30, 3 a.m., four of the astrographic plates were developed, and when dry examined. It was found that there had been a serious change of focus, so that, while the stars were shown, the definition was spoilt. This change of focus can only be attributed to the unequal expansion of the mirror through the sun's heat. The readings of the focussing scale were checked next day, but were found unaltered at 11·0 mm. It seems doubtful whether much can be got from these plates."

24. In the following table is given the value of α, the constant of the gravitational displacement, as calculated from the measures; the apparent difference of scale e between the eclipse and comparison plates; d the difference of orientation of the plates given by the measures of y, and depending on the adjustment of the plates in the measuring machine.

TABLE IX.

$(1^{r} = 12''{\cdot}3)$.

No. of Eclipse Plate.	Ref. No. of Comparison Plate.	No. of Stars.	Values of d, e, α in Revolutions at 50′ Distance.			α at Sun's Limb in Arc.
			d.	e.	α.	
			r	r	r	$''$
1	18_4	7	+0·051	+0·089	+0·033	+1·28
2	18_4	11	−0·009	+0·059	+0·025	+0·97
3	18_4	8	−0·074	+0·101	+0·028	+1·09
4	18_4	11	−0·168	+0·091	+0·033	+1·28
5	11_3	10	+0·094	+0·076	+0·025	+0·97
6	11_3	11	+0·186	+0·082	+0·021	+0·82
7	14_3	12	+0·006	+0·119	0·000	0·00
7	18_3	7	−0·054	+0·166	0·000	0·00
8	14_3	10	+0·093	+0·064	+0·021	+0·82
9	17_4	7	−0·096	+0·129	+0·008	+0·31
10	17_4	10	+0·090	+0·045	+0·026	+1·01
11	11_1	10	+0·073	+0·061	+0·032	+1·24
12	11_1	11	−0·009	+0·102	+0·049	+1·91
12	17_2	7	−0·102	+0·114	+0·019	+0·74
15	15_3	6	+0·111	+0·036	+0·018	+0·70
16	15_3	7	−0·002	+0·037	+0·018	+0·70
17	17_2	8	−0·022	+0·109	+0·012	+0·47
18	17_2	7	+0·045	0·000	+0·030	+1·17
Mean				+0·082	+0·022	+0·86

Thus the mean value of α obtained from all the astrographic plates is $0''{\cdot}86$, a figure considerably less than that obtained from the 4-inch photographs.

25. Reference to the diagram shows that the measurement of displacement depends essentially on the position of the stars Nos. 3 and 4 relative to 5 on one side and 6 and 10 on the other. These are all bright stars, and in this respect their images are

more comparable than are the images of the fainter stars. The measures of these are given in the following table :—

No. of Eclipse Plate.	Measured Values of Dy for Stars Nos.— 5	4	3	6	10	No. of Eclipse Plate.	Measured Values of Dy for Stars Nos 5	4	3	6	
	r	r	r	r	r		r	r	r	r	
1	−0·051	+0·175	+0·169	+0·201	+0·235	9	−0·059	+0·121	+0·109	+0·205	+0
2	+0·558	+0·656	+0·724	+0·668	+0·702	10	+0·033	+0·270	+0·188	+0·258	+0
3	+0·121	+0·285	+0·286	+0·274	+0·355	11	+0·025	+0·215	+0·210	+0·233	+0
4	+0·111	+0·222	+0·247	+0·231	+0·167	12	−0·068	+0·144	+0·124	+0·160	+0
5	+0·034	+0·228	+0·232	+0·218	+0·308	15	−0·038	+0·138	+0·107	+0·172	
6	+0·164	+0·488	+0·478	+0·557	+0·637	16	0·050	+0·076	+0·046	+0·127	+0
7	−0·051	+0·156	+0·162	+0·250	+0·279	17	−0·071	+0·104	+0·081	+0·186	+0
8	+0·108	+0·330	+0·314	+0·376	+0·397	18	+0·016	+0·092	+0·109	+0·099	+0

The equations given by these stars are

$$
\begin{array}{lr}
-\,0\cdot160d - 1\cdot107e - 0\cdot789a + f = Dy_5 & (1) \\
+\,0\cdot334d + 0\cdot472e + 1\cdot336a + f = Dy_4 & (2) \\
+\,0\cdot348d + 0\cdot360e + 1\cdot574a + f = Dy_3 & (3) \\
+\,0\cdot587d + 1\cdot099e + 0\cdot726a + f = Dy_6 & (4) \\
+\,0\cdot860d + 1\cdot321e + 0\cdot589a + f = Dy_{10} & (5)
\end{array}
$$

The mean of (4) and (5) added to (1) gives

$$+\,0\cdot564d + 0\cdot103e - 0\cdot131a + 2f = Dy_5 + \tfrac{1}{2}(Dy_6 + Dy_{10}).$$

While the sum of (2) and (3) gives

$$+\,0\cdot682d + 0\cdot832e + 2\cdot910a + 2f = Dy_3 + Dy_4.$$

Subtracting these we get

$$3\cdot041a + 0\cdot729e + 0\cdot118d = Dy_3 + Dy_4 - Dy_5 - \tfrac{1}{2}(Dy_6 + Dy_{10}).$$

This equation has a small coefficient for e and a very small one for d.

Calculating the quantities on the right-hand side, assuming e to be the sam all the plates, and substituting the values of d from the previous table, we find :—

$$
\begin{array}{lrclr}
 & & & \text{r} & \\
a + 0\cdot240e = + 0\cdot056 & 1 & \qquad & a + 0\cdot240e = + 0\cdot035 & 9 \\
a + 0\cdot240e = + 0\cdot049 & 2 & & a + 0\cdot240e = + 0\cdot048 & 10 \\
a + 0\cdot240e = + 0\cdot047 & 3 & & a + 0\cdot240e = + 0\cdot045 & 11 \\
a + 0\cdot240e = + 0\cdot059 & 4 & & a + 0\cdot240e = + 0\cdot059 & 12 \\
a + 0\cdot240e = + 0\cdot050 & 5 & & a + 0\cdot283e = + 0\cdot026 & 15 \\
a + 0\cdot240e = + 0\cdot059 & 6 & & a + 0\cdot240e = + 0\cdot024 & 16 \\
a + 0\cdot240e = + 0\cdot036 & 7 & & a + 0\cdot240e = + 0\cdot028 & 17 \\
a + 0\cdot240e = + 0\cdot046 & 8 & & a + 0\cdot240e = + 0\cdot029 & 18
\end{array}
$$

In photograph No. 15, star 10 is not shown, and the equation is slightly modified. It may also be noticed that the values are somewhat smaller for Nos. 15 to 18.

The means of the 16 photographs treated in this manner give

$$a + 243e = + 0^r{\cdot}0435,$$

or with the value of the scale $0^r{\cdot}082$ from the previous table

$$a = + 0^r{\cdot}024 = 0''{\cdot}93 \text{ at the limb.}$$

It may be noticed that the change of scale arising from differences of refraction and aberration is $0^r{\cdot}020$. If this value of e be taken instead of $0^r{\cdot}082$ we obtain

$$a = + 0^r{\cdot}039 = + 1''{\cdot}52 \text{ at the sun's limb.}$$

The equations on p. 311 were also solved by least squares for each plate. There is a considerable range in the deduced values of a, as is to be expected when a and e are determined independently for each plate. The mean result for a is $0''{\cdot}99$, or very nearly the same as that already found.

The photographs taken with the astrographic telescope support those obtained by the "4-inch" to the extent that they show considerable outward deflection, but for the reasons already given are of much less weight.

IV. THE EXPEDITION TO PRINCIPE.

(*Observers, Prof.* A. S. EDDINGTON *and Mr.* E. T. COTTINGHAM.)

26. The expedition left Liverpool on the "Anselm" on March 8, and travelled in company with the Sobral expedition as far as Madeira. It was necessary to wait there until April 9, when the journey was continued on the "Portugal," belonging to the Companhia Nacional de Navegação. The expedition landed at the small port of S. Antonio in the Isle of Principe on April 23.

Vice-Admiral CAMPOS RODRIGUES and Dr. F. OOM of the National Observatory, Lisbon, had kindly given us introductions, and everything possible was done by those on the island for the success of the work and the comfort of the observers. We were met on board by the Acting Administrator Sr. VASCONCÉLOS, Sr. CARNEIRO, President of the Association of Planters, and Sr. GRAGEIRA, representing the Sociedade d'Agricultura Colonial, who made all necessary arrangements. The Portuguese Government dispensed with any customs examination of the baggage.

27. Principe is a small island belonging to Portugal, situated just north of the equator in the Gulf of Guinea, about 120 miles from the African coast. The extreme length and breadth are about 10 miles and 6 miles. Near the centre mountains rise to a height of 2500 feet, which generally attract heavy masses of cloud. Except for a certain amount of virgin forest, the island is covered with cocoa plantations. The

climate is very moist, but not unhealthy. The vegetation is luxuriant, and the scenery is extremely beautiful. We arrived near the end of the rainy season, but the *gravana*, a dry wind, set in about May 10, and from then onwards no rain fell except on the morning of the eclipse.

We were advised that the prospects of clear sky at the end of May were not very good, but that the best chance was on the north and west of the island. After inspecting two other sites on the property of the Sociedade d'Agricultura Colonial, we fixed on Roça Sundy, the headquarters of Sr. CARNEIRO's chief plantation. We were Sr. CARNEIRO'S guests during our whole visit, and used freely his ample resources of abour and material at Sundy. We learnt later that he had postponed a visit to Europe n order to entertain us. We were also greatly indebted to his manager at Sundy, Sr. ATALAYA, with whom we lived for five weeks; his help and attention were nvaluable. Mr. WRIGHT and Mr. LEWIS of the Cable Station kindly assisted us as nterpreters when necessary.

Sundy is situated in the north-west of the island overlooking the sea at a height of 500 feet, and as far as possible from the cloud-gathering peaks. Our telescope was erected in a small walled enclosure adjoining the house, from which the ground sloped steeply down to the sea in the direction of the sun at eclipse. On the other side it was sheltered by a building. The approximate position was latitude 1° 40′ N., longitude 29m. 32s. E.

28. The baggage was brought to Sundy on April 28 mainly by tram, but with a break of about a kilometre, where it had to be transported through the wood by native carriers. After a week spent on the preparations, we returned to S. Antonio for the week, May 6–13, as it was undesirable to unpack the mirror so early in the damp climate. On our return to Sundy the installation and adjustments were soon completed, and the first check plates were taken on May 16. Meanwhile the gravana had begun, which, although there is no rain, is generally accompanied by increased cloud. There were, however, some days of clear sky, and the nights were usually clear.

The cœlostat was mounted on a stone pier built for the purpose. The clock weight ell into a pit below the clock deep enough to allow a run of 36 minutes without rewinding. Care was taken to use a particular part of the cœlostat-sector, considered to be the most perfect, in photographing the eclipse and the check field. The telescope Oxford astrographic object-glass, see p. 295) rested on wooden V's near the two ends, the V's being supported on packing-cases; the one at the breech-end could be moved laterally to allow of different declination settings, and was marked with an approximate declination scale. A series of exposures of one second was made on a bright star to test whether there was any shake of the telescope after inserting the plate: no shake was detected even when the exposure was made immediately; but as a safeguard for the eclipse photographs a full second was allowed to elapse before beginning the exposure. The exposure was made by moving a cardboard screen unconnected with the instrument. The telescope pointed slightly downwards, and the tube was turned so as to give the right orientation to the plate, the lines of declination being two or three degrees inclined to the horizontal. A canvas screen was arranged to protect the tube and object-glass from the direct radiation of the sun.

The adjustments call for little comment. In view of the purpose of the observations, it was desirable to adjust the tilt of the object-glass and plate with special care. It was also important that the setting on the field should be nearly exact. The sun appeared on the eclipse day in sufficient time to allow of the setting being made by means of the solar image; but arrangements had been tested by which the correct field would have been obtained if it had been cloudy up to totality.* The telescope was focussed by trial photographs of stars, and owing to the uniform temperature of the island the focus was unchanged for day observations.

The object-glass was stopped down to 8 inches for the eclipse photographs and for all check and comparison photographs used in the reductions.

29. The days preceding the eclipse were very cloudy. On the morning of May 29 there was a very heavy thunderstorm from about 10 a.m. to 11.30 a.m.—a remarkable occurrence at that time of year. The sun then appeared for a few minutes, but the clouds gathered again. About half-an-hour before totality the crescent sun was glimpsed occasionally, and by 1.55 it could be seen continuously through drifting cloud. The calculated time of totality was from 2h. 13m. 5s. to 2h. 18m. 7s. G.M.T. Exposures were made according to the prepared programme, and 16 plates were obtained. Mr. COTTINGHAM gave the exposures and attended to the driving mechanism, and Prof. EDDINGTON changed the dark slides. It appears from the results that the cloud must have thinned considerably during the last third of totality, and some star images were shown on the later plates. The cloudier plates give very fine photographs of a remarkable prominence which was on the limb of the sun.

A few minutes after totality the sun was in a perfectly clear sky, but the clearance did not last long. It seems likely that the break-up of the clouds was due to the eclipse itself, as it was noticed that the sky usually cleared at sunset.

It had been intended to complete all the measurements of the photographs on the spot; but owing to a strike of the steamship company it was necessary to return by the first boat, if we were not to be marooned on the island for several months. By the intervention of the Administrator berths, commandeered by the Portuguese Government, were secured for us on the crowded steamer. We left Principe on June 12, and after transhipping at Lisbon, reached Liverpool on July 14.

30. The following is a list of the photographs, including the comparison photographs kindly taken for us by Mr. F. A. BELLAMY at Oxford, before the instrument was dismounted. All the eclipse photographs are given, though only W and X furnished

---

* The method depended on setting the cross-wires of the theodolite (attached to the cœlostat) on a terrestrial mark, and then starting the clock at a particular instant.

results. Of the other series, only the exposures actually used in the reductions are given.

List of Plates.

Check Field (R.A. 14h. 12m. 47s., Declination +20° 30′)

Ref.	Place.	Date.	Loc. Sid. T.	Exp.	Approx. Z.D.	Bar.	Ther.	Plate.
		1919.	h. m. s.	s.	°	m.	°	
a_1	Oxford	January 16	12 55 10	60	35	29·64	37·0	S.
b_1	,,	January 17	13 10 40	60	34	29·83	35·3	S.
c_1	,,	,,	13 54 55	60	31	29·83	35·3	S.
d_1	,,	,,	14 9 25	60	31	29·83	35·3	S.
e_1	,,	January 23	13 13 30	60	33	30·45	29·0	S.
			G.M.T.					
q_1	Principe	May 22	12 25 40	40	43	29·45	76·5	S.R.
r_1	,,	,,	12 31 20	40	45	29·45	76·5	S.R.
s_2	,,	,,	12 37 50	80	46	29·45	76·5	S.R.
v_1	,,	May 25	12 22 20	40	45	29·45	76·5	S.S.
w_1	,,	,,	12 26 20	40	46	29·45	76·5	S.S.

Notes.

Column 1.—The letter is marked on the original plates (preserved at Cambridge Observatory). The number refers to the exposure, disregarding exposures taken without the 8-inch stop.

Column 2.—The co-ordinates of Oxford Observatory are 5m. 3s. W., 51° 46′ N., and of the site at Principe, 29m. 32s. E., 1° 40′ N.

Column 4.—The mid-instant of the exposure is given. Times for check plates at Principe were only noted roughly. Times for the eclipse plates are deduced from the calculated time of totality, the interval from the end of one exposure to the beginning of the next being assumed uniform.

Column 7.—Readings at Principe were taken with an aneroid recording instrument, and therefore automatically reduced to the latitude of England. The barometer during our visit was practically constant except for a regular semi-diurnal wave of amplitude about 0·05 in.

Column 9.—Brand of Plate : S. = Imperial Sovereign, S.S. = Imperial Special Sensitive, S.R. = Ilford Special Rapid, E. = Ilford Empress. Backed plates were used at Principe.

Eclipse Field (R.A. 4h. 19m. 30s., Declination +21° 43′)

Ref.	Place.	Date.	Loc. Sid. T.	Exp.	Approx. Z.D.	Bar.	Ther.	Plate.
		1919.	h. m. s.	s.	°	m.	°	
D_1	Oxford	January 16	3 58 1	5	30	29·65	39·0	S.
G_1	„	January 22	4 4 39	5	30	30·30	31·0	S.
H_1	„	„	4 34 28	5	30	30·30	31·0	S.
I_2	„	„	4 48 46	10	31	30·30	31·0	S.
K_2	„	February 9	4 45 24	10	30	30·48	24·5	S.
			G.M.T.					
K′	Principe	May 29	2 13 9	5	46	29·45	77·0	S.R.
L	„	„	2 13 28	10	46	29·45	77·0	S.R.
M	„	„	2 13 46	3	46	29·45	77·0	S.R.
N	„	„	2 14 1	5	46	29·45	77·0	E.
O	„	„	2 14 20	10	46	29·45	77·0	S.S.
P	„	„	2 14 44	15	46	29·45	77·0	S.S.
Q	„	„	2 15 6	5	46	29·45	77·0	S.R.
R	„	„	2 15 30	20	46	29·45	77·0	S.R.
S	„	„	2 15 53	3	46	29·45	77·0	S.S.
T	„	„	2 16 13	15	46	29·45	77·0	E.
U	„	„	2 16 37	10	46	29·45	77·0	S.R.
V	„	„	2 16 56	5	46	29·45	77·0	S.S.
W	„	„	2 17 15	10	46	29·45	77·0	S.
X	„	„	2 17 33	3	46	29·45	77·0	S.R.
Y	„	„	2 17 47	2	46	29·45	77·0	S.R.
Z	„	„	2 18 1	2	46	29·45	77·0	S.R.

NOTES.

Columns 1 to 9. See previous page.

The large proportion of Ilford Special Rapid plates used at the eclipse was due to the fact that experience in developing the check plates showed that these suffered less than the others from the high temperature of the water (78° F.). Ice was generally available for the check plates through the kindness of Sr. GRAGEIRA; but the supply failed after the eclipse, and formalin was used to harden the films. This was unsatisfactory except for the I.S.R. plates, and so plates P, S, T, W were brought home undeveloped. The developing at Principe was done at night, and the drying was accelerated by use of alcohol.

The use of an 8-inch stop in front of the object-glass was suggested to us by Mr. DAVIDSON, who showed that a great improvement of the images resulted; it was originally intended, however, to use the full aperture for part of totality. Early measures of check plates made at Principe soon convinced us that the results from the full aperture were greatly inferior, and we decided to rely entirely on the 8-inch aperture.

The Check Plates.

31. In addition to the eclipse field, a check field was photographed both at Oxford and at Principe. The field chosen included Arcturus, so that it was easily found with the cœlostat. Its declination was nearly the same as that of the eclipse field, and it was photographed at the same altitude at Principe in order that any systematic error, due to imperfections of the cœlostat mirror or other causes, might affect both sets of plates equally. The primary purpose was thus to check the possibility of systematic error arising from the different conditions of observation at Oxford and Principe, and from possible changes in the object-glass during transit. Unlike the Sobral expedition, we were not able to take comparison photographs of the eclipse field at Principe, because for us the eclipse occurred in the afternoon, and it would be many months before the field could be photographed in the same position in the sky before dawn. The check plates were therefore specially important for us.

As events turned out the check plates were important for another purpose, viz., to determine the difference of scale at Oxford and Principe. As shown in the report of the Sobral expedition, it is not necessary to know the scale of the eclipse photographs, since the reductions can be arranged so as to eliminate the unknown scale. If, however, a trustworthy scale is known and used in the reductions, the equations for the deflection have considerably greater weight, and the result depends on the measurement of a larger displacement. On surveying the meagre material which the clouds permitted us to obtain, it was evident that we must adopt the latter course; and accordingly the first step was to obtain from the check plates a determination of the scale of the Principe photographs.

32. All the measures were made by Prof. EDDINGTON with the Cambridge measuring machine.[*] An Oxford and a Principe plate were placed film to film so that the images of corresponding stars nearly coincided—this was possible because the Oxford plates were taken direct, and the Principe plates by reflection in the cœlostat mirror.

The small differences Δx and Δy, in the sense Principe − Oxford, were then measured for each star. Eight settings were made on each image; for half of them the field was rotated through 180 degrees by the reversion prism. Five pairs of plates were measured, and the measures are given in Table XI.

---

[*] 'Monthly Notices, R.A.S.,' vol. LXI, p. 444.

TABLE XI.—Check Plates, Measures.

Star.	Approx. Co-ords.		q_1-a_1.		w_1-b_1.		s_2-c_1.		r_1-d_1.		v_1-e_1.	
	x.	y.	Δx.	Δy.	Δx.	Δy.	Δx.	Δy.	Δx.	Δy.	Δx.	Δy.
1	1·41	20·31	4346	7180	3199	4259	6012	7375	3921	8796	5135	4399
2	5·89	12·74	3865	6405	3394	4129	4922	6132	3039	7440	5978	4170
4	9·46	11·13	3640	5932	3408	4118	4369	5366	2638	6776	5966	4441
5	12·00	6·84	3311	5590	—	—	3831	4752	1938	6156	—	4314
6	12·80	27·33	5415	6561	3192	5140	7689	5925	5379	7580	5032	5794
7	13·75	13·78	4076	5630	3496	4290	4891	4805	3101	6461	5906	4826
8	15·50	24·38	5125	6300	—	—	—	—	—	—	5139	5412
10	20·13	10·49	3965	4940	3679	4505	4656	3568	2866	5370	6398	5229
11	20·81	0·93	2874	4352	3876	3759	2845	2815	1238	4758	7268	4482
12	22·91	6·23	3685	4436	3931	4158	4039	2738	2270	4551	6765	5076
13	26·46	8·96	4222	4288	4045	4326	4724	2232	2720	4120	6836	5561

The unit for x and y is 5 millimetres, which is approximately equal to 5′. The differences Δx, Δy are given in units of the fifth place of decimals $= 0''{\cdot}003$. The centre of the plate is near $x = 14$, $y = 14$.

Plate-constants were then calculated in the usual way, by the formulæ

$$\Delta x = ax + by + c$$

$$\Delta y = dx + ey + f$$

These were applied, and the residuals $\Delta_1 x$, $\Delta_1 y$ converted into arc are as follows :—

TABLE XII.—Check Plates, Residuals.

Star.	q_1-a_1.		w_1-b_1.		s_2-c_1.		r_1-d_1.		v_1-e_1.		Mean.	
	$\Delta_1 x$.	$\Delta_1 y$.	$\Delta_1 x$.	$\Delta_1 y$.	$\Delta_1 x$.	$\Delta_1 y$.	$\Delta_1 x$.	$\Delta_1 y$.	$\Delta_1 x$.	$\Delta_1 y$.	$\Delta_1 x$.	$\Delta_1 y$.
	″	″	″	″	″	″	″	″	″	″	″	″
1	−0·02	−0·02	+0·29	−0·34	+0·02	−0·07	−0·03	+0·22	+0·49	+0·01	+0·15	−0·04
2	+0·39	+0·15	+0·16	+0·14	+0·69	0·00	+0·69	−0·29	+0·10	−0·23	+0·41	−0·05
4	−0·14	−0·04	−0·16	+0·09	−0·38	−0·12	−0·02	−0·37	−0·54	+0·12	−0·25	−0·06
5	−0·08	+0·35	—	—	+0·25	+0·19	−0·21	−0·21	—	−0·01	−0·01	+0·08
6	−0·06	−0·10	−0·28	+0·27	−0·09	+0·14	−0·10	+0·12	+0·15	+0·49	−0·08	+0·18
7	−0·06	−0·28	−0·10	−0·16	−0·74	−0·09	−0·31	+0·02	−0·39	−0·12	−0·32	−0·13
8	−0·30	+0·34	—	—	—	—	—	—	−0·38	−0·68	−0·34	−0·17
10	−0·02	−0·10	−0·21	+0·52	−0·15	+0·16	+0·08	+0·25	−0·08	+0·34	−0·08	+0·23
11	−0·46	−0·01	−0·13	−0·22	−0·13	+0·11	−0·13	+0·71	+0·30	−0·28	−0·11	+0·06
12	+0·16	−0·14	+0·13	−0·04	+0·19	−0·06	+0·17	−0·09	−0·13	+0·08	+0·10	−0·05
13	+0·59	−0·12	+0·32	−0·26	+0·34	−0·25	−0·13	−0·38	+0·48	+0·28	+0·32	−0·15

The mean residual without regard to sign is $\pm 0''{\cdot}21$, from which the probable error of a determination of Δx or Δy is $\pm 0''{\cdot}22$.

Star 7 is much the brightest. Stars 1, 6, 11, 13 are rather bright. Stars 2, 4, 10, 12 are fainter and more comfortable to measure. Stars 5 and 8 are very faint. Arcturus is on the plates but is much too bright to measure. No measures have been rejected.

The determination of the deflection on the eclipse plates is based on the declinations (y), and the last column of Table XII. shows that on the check plates the y-comparisons are free from any serious systematic error.

Star 7 is of particular interest; its position near the centre of the field corresponds to that of κ_1, κ_2 Tauri in the eclipse field, from which the greatest deflection is expected. The images (which are not quite round) have the same characteristic shape. Further, the brightness of No. 7 corresponds with, but exaggerates, the brightness of κ_1 Tauri which is the brightest star in the eclipse field. It is therefore a valuable check to find that its systematic error in declination is insignificant compared with the displacement (of the order of 1″) afterwards found for κ_1 and κ_2 Tauri.

The systematic errors in right ascension are larger (probably through imperfect driving of the clock). They may affect the displacement indirectly through the orientation constant, but with much reduced effect. Allowing for this reduction in importance there appears to be nothing to trouble about.

The primary purpose of the check plates is thus fulfilled. They show that photographs of a check field of stars taken at Oxford and Principe show none of the displacements which are exhibited by the photographs of the eclipse field taken under precisely similar instrumental conditions. The inference is that the displacements in the latter case can only be attributed to presence of the eclipsed sun in the field.

33. We turn now to the differences of scale between Oxford and Principe, which are given by the plate-constants a, b, d, e determined from the measures. As determined, these include the effects of differential refraction and aberration. The latter corrections were calculated for each plate by the usual formulæ and applied, so as to determine the corrected plate-constants, a', b', d', e' free from differential refraction and aberration. Due allowance was made for the change in the coefficient of refraction owing to the difference of barometer and temperature (about 40°) between Oxford and Principe. The results are as follows (in units of the fifth place of decimals):—

TABLE XIII.—Check Plates, Plate-Constants.

Comparison.	Uncorrected.				Corrected.				
	a.	*b.*	*d.*	*e.*	*a′.*	*b′.*	*d′.*	*e′.*	*b′+d′.*
$q_1 - a_1$	+32·7	+101·0	− 87·8	+58·2	+32·7	+ 98·4	− 90·4	+32·1	+ 8·0
$w_1 - b_1$	+26·2	− 16·0	+ 25·9	+53·6	+30·4	− 22·5	+ 19·4	+31·4	− 3·1
$s_2 - c_1$	+31·5	+192·5	−173·5	+64·8	+35·8	+182·6	−183·4	+42·1	− 0·8
$r_1 - d_1$	+28·2	+165·0	−146·8	+69·8	+32·1	+157·8	−154·0	+45·0	+ 3·8
$v_1 - e_1$	+21·6	− 76·2	+ 70·6	+61·4	+25·2	− 80·5	+ 66·3	+35·7	−14·2
				Mean	+31·2	—	—	+37·3	− 1·3

The sign of the results shows that the scale of the photographs is larger at Principe than at Oxford; in fact the focus must have been set about 1·2 mm. further out (apart from any change of length compensated by expansion of the photographic plates). As the error in focussing was probably not more than 0·5 mm., the greater part of this shift must be due to the focal length of the lens combination increasing with temperature more rapidly than the linear expansion of the glass.

If the only difference were a change of focal length, we should have $a' = e'$. There is a fairly strong indication that e' is greater than a'. This is no doubt due to a change in the definition caused by the cœlostat mirror or by a shift of the object-glass lenses on the journey; and, as it will presumably affect the eclipse plates in the same way, it is best to adopt the values of a' and e' as determined, rather than to take a mean. In so doing we shall at any rate not exaggerate the displacement, which depends mainly on the y-measures and is reduced by adopting too large a value of e'.*

The difference $b' - d'$ merely gives the relative orientation of the two plates as placed face to face. The sum $b' + d'$ practically vanishes, as it should do. However, for consistency we adopt the small value found.

From the internal discordances of our determination of e' (the most important of these constants) the probable error of the mean is $\pm 2\cdot 1$. This, as shown later, will cause a probable error of our final determination of the deflection, reduced to the limb of the sun, of amount $\pm 0''\cdot 14$, affecting all determinations systematically. Errors in the other constants have much smaller influence.

The Eclipse Plates.

34. The eclipse plates from K to S show no star images. After that the cloud lightened somewhat, and some images appear on the remaining plates. The sky was never clear and nothing fainter than $5'\cdot 5$ is shown. The cloud was variable in different parts of the plate, so that the brightness of the images varies erratically and the diffusion is also variable.

In order to obtain results of any weight the stars 4 and 3 (κ_1 and κ_2 Tauri), which theoretically should be strongly displaced, must be shown. They appear on all plates from T to Z, and being near the centre of the field have good images. They are relatively rather faint on plate U, but are bright on the other plates. The appearance of the remaining stars is as follows:—

Plate T.	6 bright; 10 faint.
Plate U.	6, 10 very bright; 11 faint.
Plate V.	6 bright; 10 fair.
Plate W.	5, 6 good; 10 diffused.
Plate X.	5, 6, 11 good.
Plate Y.	5, 6, 11 faint, diffused; 12 very faint.
Plate Z.	5, 6, 11 faint, diffused.

---

* It happens that it is also reduced, but to a less extent, by using too small a value of a'.

The possibility of a determination of deflection practically depends on the appearance of star 5. The relative displacement of 5 and 3 is on EINSTEIN'S theory, $1''\cdot 2$ in the y-co-ordinate. Further, the x-measures of 5 are needed for a really good determination of the orientation. Star 11 can scarcely take its place. It is true that the relative displacement is then $0''\cdot 8$; but the orientation affects this with a much larger factor, and the orientation is badly determined in the absence of star 5.

Accordingly plates W and X are the only ones likely to give a trustworthy result. X is somewhat the better plate of the two.* Measures have been made of the faint diffused images on plates Y and Z; but, as might have been expected, they are hopelessly discordant and cannot be reconciled by any adopted value of the deflection.

35. We give the measures of plates X and W in detail. Both comparisons of X were measured at Principe a few days after the eclipse. Plate W, which was not developed until after the return of the expedition, was measured at Cambridge on August 22–23.†

Plate X.

(1) Comparison with Oxford Plate G_1.

The differential refraction for all the eclipse plates is

$$a = -46\cdot 5, \qquad b, d = +8\cdot 2, \qquad e = -27\cdot 0$$

the differential aberration being zero.

For the comparison plate G_1

$$a = -19\cdot 1, \qquad b, d = +0\cdot 7, \qquad e = -28\cdot 3.$$

Hence for $X - G_1$

$$a = -27\cdot 4, \qquad b, d = +7\cdot 5, \qquad e = +1\cdot 3.$$

---

* Plate X has also the merit of a short exposure, 3s. We should mistrust the x-measures of a long exposure with variable cloud and imperfect guiding, because there is nothing to show that the images of the different stars are formed at the same time.

† Of the comparisons of check plates, $w_1 - b_1$ was measured on August 20, and the others about the end of September. Previous measures had been made at Principe with three earlier check plates taken on the night of May 16; but a slight change of adjustment of tilt was made the following day (thereafter it remained unaltered until the eclipse), and the small change of focus allowed for in the comparisons. These furnished a provisional scale which was used to obtain preliminary results. Afterwards the measurement of check plates was undertaken in a more systematic way, using later plates about which no doubt could arise, and giving the results printed above. No change of any importance was found; the final value for the deflection at the limb was reduced by $0''\cdot 4$ compared with the provisional value, but this was mainly due to the adoption of separate values of a' and e' instead of adopting the mean, and to recalculation of the differential refraction and aberration.

To these must be added the terms representing change of scale, determined from the check plates (Table XIII.), viz.,

$$a = +31 \cdot 2, \qquad b, d = -0 \cdot 6, \qquad e = +37 \cdot 3.$$

Hence the whole difference X — G_1 is given by

$$a = +3 \cdot 8, \qquad b, d = +6 \cdot 9, \qquad e = +38 \cdot 6.$$

The first step is to take the measured differences Δx, Δy, and take out the parts $ax + by$, $dx + ey$, due to these terms, leaving the corrected differences $\Delta_1 x$, $\Delta_1 y$.

$\Delta_1 x$ and $\Delta_1 y$ contain (1) the Einstein displacement, if any, and (2) the unknown relative orientation of the plates giving rise to terms of the form, $\Delta x = +\theta y$, $\Delta y = -\theta x$. These two parts could be separated by a least-squares solution, but in view of the poor quality of the material it seems better to adopt a method which keeps a better check on possible discordances and shows more clearly what is happening. The Einstein displacement in x is small, and we might perhaps neglect it altogether in determining θ from the x-measures. However, it is clear from preliminary trials that a displacement exists—whether the half or the full Einstein displacement. Hence if we take out three-quarters of the full Einstein displacement ($\frac{3}{4}E_x$) we divide the already slight effect by 4, and at the same time deal fairly between the two hypotheses.* The residuals $\Delta_2 x$ result.

From the equations $\Delta_2 x = c + \theta y$ we determine by least squares the orientation θ, which is found to be $+163$. Removing the term $163y$ we obtain the residuals $\Delta_3 x$.

Turning to $\Delta_1 y$, we correct for the orientation by taking out the term $-163x$, leaving $\Delta_3 y$. These values should agree for all the stars, except for the displacement and the accidental error.

Denoting the value of the displacement at 50′ (or 10 réseau-intervals) from the centre of the sun by κ, the y-displacements of the various stars will be $\kappa\alpha_y$, where α_y has the values tabulated below. We can therefore obtain κ by solving by least-squares the equations

$$\Delta_3 y = f + \kappa\alpha_y.$$

The radius of the sun during the eclipse was $15' \cdot 78$. Hence the full Einstein displacement of $1'' \cdot 75$ corresponds to $0'' \cdot 55$ at 50′ distance, or, in our units of $0'' \cdot 003$, $\kappa = 184$. It is easily seen that the value is somewhere near this, and it is therefore easier and more instructive to take out $E_y = 184\alpha_y$, and determine the correction to κ from the residuals $\Delta_4 y$. We also remove the mean of $\Delta_4 y$ obtaining the final residuals.

The normal equations corresponding to equations of condition

$$\text{residual} = \delta f + \alpha_y \, \delta\kappa$$

---

* The smaller the displacement provisionally assumed for x, the larger is the displacement ultimately found from y (see p. 327).

are found to be

$$5\delta f + 2\cdot83\delta\kappa = -\ 1$$

$$2\cdot83\delta f + 4\cdot83\delta\kappa = +64$$

whence

$$3\cdot23\delta\kappa = +64,$$

$$\delta\kappa = +20.$$

An increase of 20 on 184 corresponds to an increase of $0''\cdot19$ on $1''\cdot75$. Hence the resulting deflection at the limb is $1''\cdot94$.

Since the full deflection is indicated we complete the results for x by taking out the remaining $\frac{1}{4}E_x$, obtaining $\Delta_4 x$, and then tabulate the residuals from the mean values −5942.

The successive steps are shown below :—

Star.	x.	Δx.	$3\cdot8x$.	$6\cdot9y$.	$\Delta_1 x$.	$\frac{3}{4}E_x$.	$\Delta_2 x$.	$+163y$.	$\Delta_3 x$.	$\Delta_4 x$.	Resid. (unit = $0''\cdot003$).
11	1·39	−3916	5	86	−4007	−76	−3931	2021	−5952	−5927	+ 15
5	12·40	−5518	47	20	−5585	−79	−5506	478	−5984	−5958	− 16
4	17·34	−2869	66	129	−3064	−54	−3010	3051	−6061	−6043	−101
3	17·48	−2924	66	121	−3111	−69	−3042	2869	−5911	−5888	+ 54
6	19·87	−1568	75	172	−1815	+ 3	−1818	4075	−5893	−5894	+ 48

Star.	y.	Δy.	$6\cdot9x$.	$38\cdot6y$.	$\Delta_1 y$.	$-163x$.	$\Delta_3 y$.	E_y.	$\Delta_4 y$.	α_y.	Resid.
11	12·40	6398	10	479	5909	− 227	6136	+ 6	6130	+0·03	+ 5
5	2·93	4121	86	113	3922	−2021	5943	−127	6070	−0·69	− 55
4	18·72	4512	120	722	3670	−2826	6496	+234	6262	+1·27	+137
3	17·60	4236	121	679	3436	−2849	6285	+272	6013	+1·48	−112
6	24·99	4148	137	965	3046	−3239	6285	+136	6149	+0·74	+ 24

(2) Comparison with Oxford Plate H_1.

The reductions are similar and are given in a rather more condensed form below. The theoretical plate constants are

$$a = +3\cdot8, \qquad b, d = +8\cdot3, \qquad e = +38\cdot6.$$

Star.	Δx.	$\Delta_1 x$.	$\Delta_2 x$.	$+10y$.	$\Delta_3 x$.	$\Delta_4 x$.	Resid.
11	7290	7182	7258	124	7134	7159	+235
5	6751	6680	6759	29	6730	6756	−168
4	7126	6905	6959	187	6772	6790	−134
3	7320	7108	7177	176	7001	7024	+100
6	7429	7147	7144	250	6894	6893	− 31

Star.	Δy.	$\Delta_1 y$.	$-10x$.	$\Delta_3 y$.	E_y.	$\Delta_4 y$.	Resid.
11	1586	1095	− 14	1109	+ 6	1103	+172
5	858	642	−124	766	−127	893	− 38
4	1881	1015	−173	1188	+234	954	+ 23
3	1785	961	−175	1136	+272	864	− 67
6	1909	779	−199	978	+136	842	− 89

The normal equations are

$$5\delta f + 2\cdot83\delta\kappa = +1$$

$$2\cdot83\delta f + 4\cdot83\delta\kappa = -105$$

whence

$$3\cdot23\delta\kappa = -105,$$

$$\delta\kappa = -33.$$

The corresponding deflection at the limb is

$$1''\cdot75 - 0''\cdot31 = 1''\cdot44.$$

Plate W.

Although the exposure was only 10 seconds the images have jumped in R.A., so that the appearance is dumb-bell shaped. They are, however, symmetrical, so that fair measures of x can be made; the y measures on which the result chiefly depends are unaffected. Star 10 is very diffused in R.A.

(1) Comparison with Oxford Plate D_1.

Theoretical plate-constants

$$a = +4\cdot9, \qquad b, d = +6\cdot5, \qquad e = +39\cdot7.$$

Star.	x.	Δx.	$\Delta_1 x$.	$\frac{3}{4}E_x$.	$\Delta_2 x$.	$+91y$.	$\Delta_3 x$.	$\Delta_4 x$.	Resid.
5	12·40	2450	2370	−79	2449	267	2182	2208	+ 40
4	17·34	3948	3741	−54	3795	1704	2091	2109	− 59
3	17·48	3834	3634	−69	3703	1602	2101	2124	− 44
6	19·87	4525	4266	+ 3	4263	2275	1988	1987	−181
10	22·60	5199	4911	+17	4894	2476	2418	2412	+244

Star.	y.	Δy.	$\Delta_1 y$.	$-91x$.	$\Delta_3 y$.	E_y.	$\Delta_4 y$.	α_y.	Resid.
5	2·93	5320	5123	−1128	6251	−127	6378	−0·69	+ 70
4	18·72	5745	4889	−1578	6467	+234	6233	+1·27	− 75
3	17·60	5911	5098	−1591	6689	+272	6417	+1·48	+109
6	24·99	5628	4507	−1808	6315	+136	6179	+0·74	−129
10	27·21	5616	4389	−2057	6446	+114	6332	+0·62	+ 24

Normal equations

$$5\delta f + 3\cdot42\delta\kappa = -1$$

$$3\cdot42\delta f + 5\cdot21\delta\kappa = -62$$

whence

$$2\cdot87\delta\kappa = -61$$

$$\delta\kappa = -21.$$

Hence deflection at the limb is

$$1''\cdot75 - 0''\cdot20 = 1''\cdot55.$$

(2) Comparison with Oxford Plate I_2.
Theoretical plate constants

$$a = +4\cdot0, \qquad b, d = +9\cdot1, \qquad e = +38\cdot8.$$

Star.	Δx.	$\Delta_1 x$.	$\Delta_2 x$.	$-30y$.	$\Delta_3 x$.	$\Delta_4 x$.	Resid.
5	5050	4973	5052	− 88	5140	5166	+ 46
4	4732	4493	4547	−562	5109	5127	+ 7
3	4622	4392	4461	−528	4989	5012	−108
6	4635	4329	4326	−750	5076	5075	− 45
10	4764	4426	4409	−816	5225	5219	+ 90

Star.	Δy.	$\Delta_1 y$.	$+30x$.	$\Delta_3 y$.	E_y.	$\Delta_4 y$.	Resid.
5	−6824	−7051	372	−7423	−127	−7296	− 15
4	−5751	−6635	520	−7155	+234	−7389	−108
3	−5609	−6451	524	−6975	+272	−7247	+ 34
6	−5425	−6576	596	−7172	+136	−7308	− 27
10	−5109	−6371	678	−7049	+114	−7163	+118

Normal equations

$$5\delta f + 3\cdot42\delta\kappa = +2$$

$$3\cdot42\delta f + 5\cdot21\delta\kappa = -24$$

whence

$$2\cdot87\delta\kappa = -25,$$

$$\delta\kappa = -9.$$

Hence deflection at the limb is

$$1''\cdot75 - 0''\cdot08 = 1''\cdot67.$$

Plate U.

Comparison with Oxford Plate K_2.

Since Plate U shows some good images it has been examined, although owing to the absence of star 8 the weight is small. The measures were made at Principe.

Theoretical plate-constants

$$a = +2\cdot 8, \qquad b, d = +8\cdot 9, \qquad e = +37\cdot 7.$$

Star.	x.	Δx.	$\Delta_1 x$.	$+240y$.	E_x.	$\Delta_4 x$.	Resid.
11	1·39	2905	2791	2976	−101	− 84	−147
4	17·34	4508	4292	4493	− 72	−129	−192
3	17·48	4626	4420	4224	− 92	+288	+225
6	19·87	6270	5992	5998	+ 4	− 10	− 73
10	22·60	7110	6805	6530	+ 23	+252	+189

Star.	x.	Δy.	$\Delta_1 y$.	$-240x$.	E_y.	$\Delta_4 y$.	Resid.
11	12·40	9026	8547	− 334	+ 6	8875	− 94
4	18·72	5846	4986	−4162	+234	8914	− 55
3	17·60	5985	5165	−4195	+272	9089	+120
6	24·99	5458	4339	−4769	+136	8972	+ 3
10	27·21	4911	3684	−5424	+114	8994	+ 25

In this case it is not possible to determine the orientation with sufficient accuracy from the x-measures; the value here applied is an arbitrary preliminary value. We accordingly make a least-squares solution from both x- and y-residuals to determine the correction to the orientation, $\delta\theta$, as well as δc, δf and $\delta\kappa$.

The result is

$$\delta\theta = +2, \qquad \delta\kappa = +121.$$

This gives the deflection

$$2''\cdot 90.$$

The probable error is, however, ± $0''\cdot 87$, so that the result is practically worthless. Further, it is much more likely to be affected by systematic error than the previous results.

The large probable error is partly due to the large residuals which are greater than in the previous measures; in particular star 3 is unduly faint. If the same accuracy had been obtained, the theoretical weight would have been half that of plates W and X;

but having regard to possible systematic error, probably a quarter weight would more nearly represent the true value.

This determination is ignored in the subsequent discussion.

36. It is easy to calculate the effects of any errors in the adopted scale, orientation, &c., on the final result (deflection at the limb). We give some illustrations.

An error in the adopted scale of y of 10 units (in the fifth place of decimals) would lead to an error 0″·68 in the result from either plate. Thus the probable error ±2·1 in the determination of e' gives a probable error ±0″·14 in the final result; or, if we adopted the largest (rather discordant) value found for e' instead of the mean, we should reduce the result by 0″·52.

An error of 10 units in the orientation gives an error in the result of 0″·45 for plate X, and 0″·22 for Plate W. It is therefore of less importance, and further it is not likely to be systematic.

Errors in the measurement of x only affect the result through the orientation. For Plate X, a probable error of ±0″·20 in the x-measures would give an error ±4·0 in the orientation, leading to an error ±0″·18 in the result; whereas an error of the same magnitude in the y measures gives directly an error ±0″·35 in the result. For Plate W, the probable error of ±0″·20 in x gives an error ±3·5 in the orientation and ±0″·08 in the result, compared with ±0″·38 for similar inaccuracy in y. It is particularly fortunate that the x-measures are so unimportant for Plate W, because, as already mentioned, the images trailed on that plate.

Finally, it will be remembered that in order not to commit ourselves to the Einstein hypothesis prematurely we neglected the correction $\frac{1}{4}\mathrm{E}_{,}$ in determining the orientation. This will make a difference of 0″·029 in the results from Plate W and 0″·092 from Plate X. The effect is that the deduced deflection needs to be decreased, and the mean correction —0″·06 should be applied to the mean result obtained, or rather, to make the adopted deflection for x consistent with the deduced value from y, the correction needed is —0″·04.

Discussion of the Results.

37. The four determinations from the two eclipse plates are

X − G	1″·94
X − H	1″·44
W − D	1″·55
W − I	1″·67

giving a mean of 1″·65.

They evidently agree with EINSTEIN'S predicted value 1″·75.

The residuals* in the separate comparisons reduced to arc are as follows. They do not appear to show any special peculiarities.

Star.	x residuals.					y residuals.				
	G.	H.	D.	I.	Mean.	G.	H.	D.	I.	Mean.
	″	″	″	″	″	″	″	″	″	″
11	+0·04	+0·70	—	—	—	+0·01	+0·52	—	—	—
5	−0·05	−0·50	+0·12	+0·14	−0·07	−0·16	−0·11	+0·21	−0·04	−0·02
4	−0·30	−0·40	−0·18	+0·02	−0·21	+0·41	+0·07	−0·22	−0·32	−0·02
3	+0·16	+0·30	−0·13	−0·32	0·00	−0·34	−0·20	+0·33	+0·10	−0·03
6	+0·14	−0·09	−0·54	−0·13	−0·16	+0·07	−0·27	−0·39	−0·08	−0·17
10	—	—	+0·73	+0·27	—	—	—	+0·07	+0·35	—

The average y-residual is ±0″·22, which gives a probable error for y of ±0″·21. It is satisfactory that this agrees so nearly with the probable error (±0″·22) of the check plates, showing that the images are of about the same degree of difficulty and therefore presumably comparable. The probable error of x is ±0″·25, but we are not so much concerned with this.

The weight of the determination of $\delta\kappa$ is about 3 (strictly 3·23 for Plate X and 2·87 for Plate W). The probable error of κ is therefore ±0″·12, which corresponds to a probable error of ±0″·38 in the final values of the deflection.

As the four determinations involve only two eclipse plates and are not wholly independent, and further small accidental errors may arise through inaccurate determination of the orientation, the probable error of our mean result will be about ±0″·25. There is further the error of ±0″·14 affecting all four results equally, arising from the determination of scale. Taking this into account, and including the small correction −0″·04 previously mentioned, our result may be written

$$1''\cdot61\pm0''\cdot30.$$

It will be seen that the error deduced in this way from the residuals is considerably larger than at first seemed likely from the accordance of the four results. Nevertheless the accuracy seems sufficient to give a fairly trustworthy confirmation of EINSTEIN'S theory, and to render the half-deflection at least very improbable.

38. It remains to consider the question of systematic error. The results obtained with a similar instrument at Sobral are considered to be largely vitiated by systematic

---

* The residuals refer to the theoretical deflection 1″·75, not the deduced deflections.

errors. What ground then have we—apart from the agreement with the far superior determination with the 4-inch lens at Sobral—for thinking that the present results are more trustworthy ?

At first sight everything is in favour of the Sobral astrographic plates. There are 12 stars shown against 5, and the images though far from perfect are probably superior to the Principe images. The multiplicity of plates is less important, since it is mainly a question of systematic error. Against this must be set the fact that the five stars shown on Plates W and X include all the most essential stars ; stars 3 and 5 give the extreme range of deflection, and there is no great gain in including extra stars which play a passive part. Further, the gain of nearly two extra magnitudes at Sobral must have meant over-exposure for the brighter stars, which happen to be the really important ones ; and this would tend to accentuate systematic errors, whilst rendering the defects of the images less easily recognised by the measurer. Perhaps, therefore, the cloud was not so unkind to us after all.

Another important difference is made by the use of the extraneous determination of scale for the Principe reductions. Granting its validity, it reduces very considerably both accidental and systematic errors. The weight of the determination from the five stars with known scale is more than 50 per cent. greater than the weight from the 12 stars with unknown scale. Its effect as regards systematic error may be seen as follows. Knowing the scale, the greatest relative deflection to be measured amounts to 1″·2 on EINSTEIN'S theory ; but if the scale is unknown and must be eliminated, this is reduced to 0·″67. As we wish to distinguish between the full deflection and the half deflection, we must take half these quantities. Evidently with poor images it is much more hopeful to look for a difference of 0″·6 than for 0″·3. It is, of course, impossible to assign any precise limit to the possible systematic error in interpretation of the images by the measurer ; but we feel fairly confident that the former figure is well outside possibility.

A check against systematic error in our discussion is provided by the check plates, as already shown. Its efficacy depends on the similarity of the images on the check plates and eclipse plates at Principe. Both sets are fainter than the Oxford images with which they are compared, the former owing to the imperfect driving of the cœlostat, which made it impossible to secure longer exposures, the latter owing to cloud. Both sets have a faint wing in declination, but this is separated by a slight gap from the true images, and, at least on the plates measured, the wing can be distinguished and ignored. The images on Plates W and X are not unduly diffused except for No. 10 on Plate W. Difference in quality between the eclipse images and the Principe check images is not noticeable, and is certainly far less than the difference between the latter and the Oxford images ; and, seeing that the latter comparison gives no systematic error in y, it seems fair to assume that the comparison of the eclipse plates is free from systematic error.

The writer must confess to a change of view with regard to the desirability of using an extraneous determination of scale. In considering the programme it had seemed too risky a proceeding, and it was thought that a self-contained determination would receive more confidence. But this opinion has been modified by the very special circumstances at Principe; and it is now difficult to see that any valid objection can be brought against the use of the scale.

The temperature at Principe was remarkably uniform and the extreme range probably did not exceed 4° during our visit—including day and night, warm season and cold season. The temperature ranged generally from 77½° to 79½° in the rainy season, and about 1° colder in the cool gravana. All the check plates and eclipse plates were taken within a degree of the same temperature, and there was, of course, no perceptible fall of temperature preceding totality. To avoid any alteration of scale in the daytime the telescope tube and object-glass were shaded from direct solar radiation by a canvas screen; but even this was scarcely necessary, for the clouds before totality provided a still more efficient screen, and the feeble rays which penetrated could not have done any mischief. A heating of the mirror by the sun's rays could scarcely have produced a true alteration of scale though it might have done harm by altering the definition; the cloud protected us from any trouble of this kind. At the Oxford end of the comparison the scale is evidently the same for both sets of plates, since they were both taken at night and intermingled as regards date.

It thus appears that the check scale is legitimately applicable to the eclipse plates. But the method may not be so satisfactory at future eclipses, since the particular circumstances at Principe are not likely to be reproduced. As regards other sources of systematic error, our chief guarantee lies in the comparatively large amount of the deflection to be measured, and the test satisfied by the check plates that photographs of another field under similar conditions show no deflections comparable with those here found.

V. GENERAL CONCLUSIONS.

39. In summarising the results of the two expeditions, the greatest weight must be attached to those obtained with the 4-inch lens at Sobral. From the superiority of the images and the larger scale of the photographs it was recognised that these would prove to be much the most trustworthy. Further, the agreement of the results derived independently from the right ascensions and declinations, and the accordance of the residuals of the individual stars (p. 308) provides a more satisfactory check on the results than was possible for the other instruments.

These plates gave

From declinations . . . . . $1''\cdot94$

From right ascensions . . . . $2''\cdot06$

The result from declinations is about twice the weight of that from right ascensions, so that the mean result is

$$1''\cdot98$$

with a probable error of about ±0″·12.

The Principe observations were generally interfered with by cloud. The unfavourable circumstances were perhaps partly compensated by the advantage of the extremely uniform temperature of the island. The deflection obtained was

$$1''\cdot61.$$

The probable error is about ±0″·30, so that the result has much less weight than the preceding.

Both of these point to the full deflection 1″·75 of EINSTEIN'S generalised relativity theory, the Sobral results definitely, and the Principe results perhaps with some uncertainty. There remain the Sobral astrographic plates which gave the deflection

$$0''\cdot93$$

discordant by an amount much beyond the limits of its accidental error. For the reasons already described at length not much weight is attached to this determination.

It has been assumed that the displacement is inversely proportional to the distance from the sun's centre, since all theories agree on this, and indeed it seems clear from considerations of dimensions that a displacement, if due to gravitation, must follow this law. From the results with the 4-inch lens, some kind of test of the law is possible though it is necessarily only rough. The evidence is summarised in the following table and diagram, which show the radial displacement of the individual stars (mean from all the plates) plotted against the reciprocal of the distance from the centre. The displacement according to EINSTEIN'S theory is indicated by the heavy line, according to the Newtonian law by the dotted line, and from these observations by the thin line.

RADIAL Displacement of Individual Stars.

Star.	Calculation.	Observation.
	″	″
11	0·32	0·20
10	0·33	0·32
6	0·40	0·56
5	0·53	0·54
4	0·75	0·84
2	0·85	0·97
3	0·88	1·02

Thus the results of the expeditions to Sobral and Principe can leave little doubt that a deflection of light takes place in the neighbourhood of the sun and that it is of the amount demanded by Einstein's generalised theory of relativity, as attributable to the sun's gravitational field. But the observation is of such interest that it will probably be considered desirable to repeat it at future eclipses. The unusually favourable conditions of the 1919 eclipse will not recur, and it will be necessary to photograph fainter stars, and these will probably be at a greater distance from the sun.

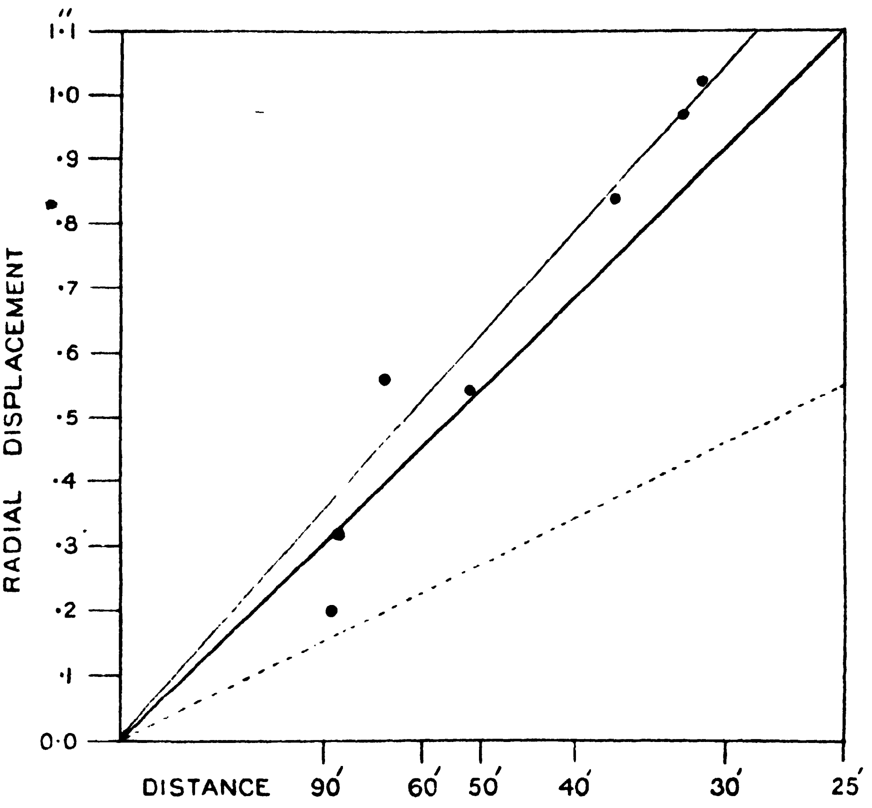
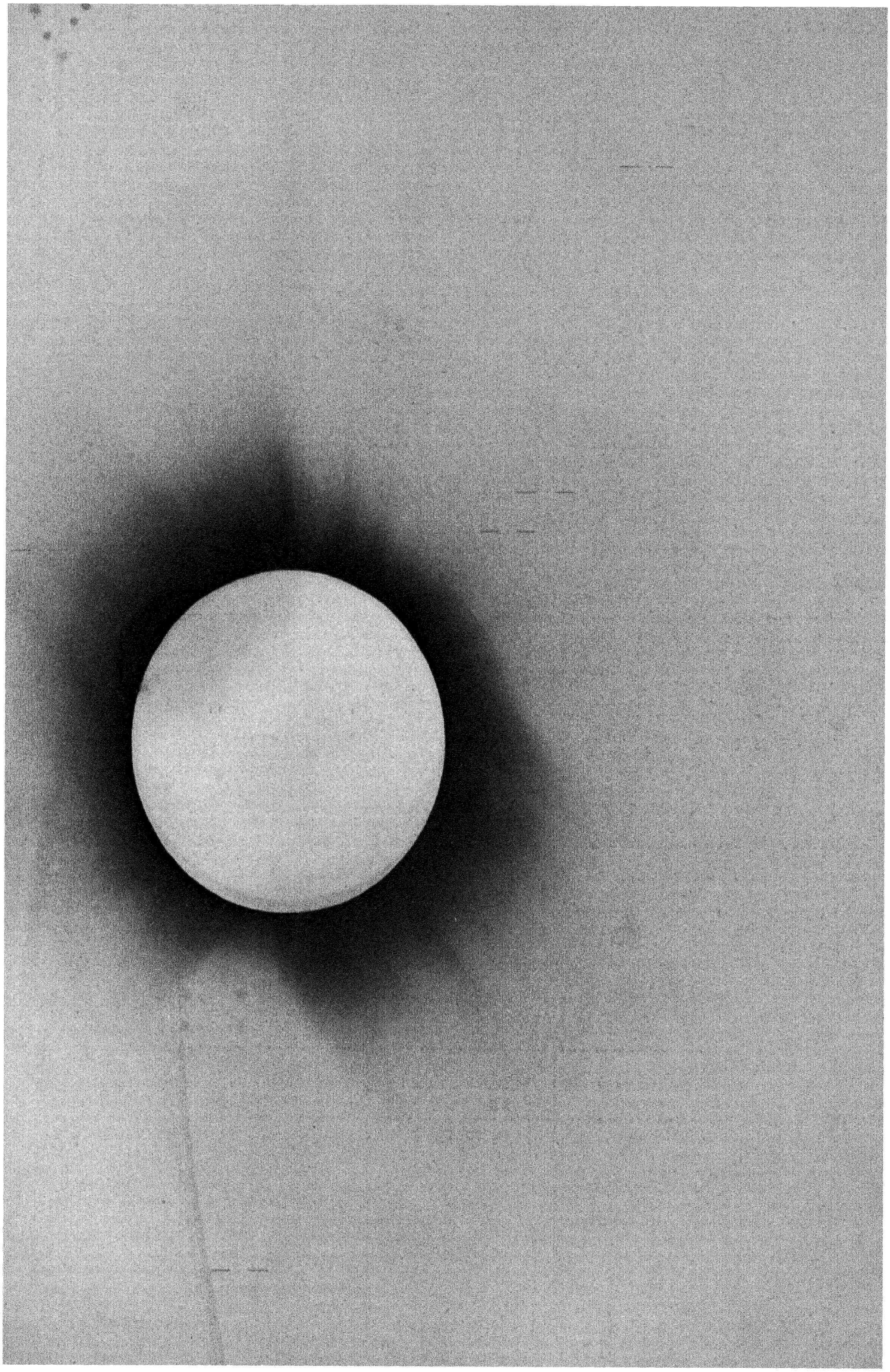


Diagram 2.

This *can* be done with such telescopes as the astrographic with the object-glass stopped down to 8 inches, if photographs of the same high quality are obtained as in regular stellar work. It will probably be best to discard the use of cœlostat mirrors. These are of great convenience for photographs of the corona and spectroscopic observations, but for work of precision of the high order required, it is undesirable to introduce complications, which can be avoided, into the optical train. It would seem that some form of equatorial mounting (such as that employed in the Eclipse Expeditions of the Lick Observatory) is desirable.

In conclusion, it is a pleasure to record the great assistance given to the Expeditions from many quarters. Reference has been made in the course of the paper to some of these. Especial thanks are due to the Brazilian Government for the hospitality and facilities accorded to the observers in Sobral. They were made guests of the



Government, who provided them with transport, accommodation and labour. Dr. MORIZE, Director of the Rio Observatory, acting on behalf of the Brazilian Government, made most complete arrangements for the Expedition, and in this way contributed materially to its success.

On behalf of the Principe Expedition, special thanks are due to Sr. JERONYMO CARNEIRO, who most hospitably entertained the observers and provided for all their requirements, and to Sr. ATALAYA, whose help and friendship were of the greatest service to the observers in their isolated station.

We gratefully acknowledge the loan for more than six months of the astrographic object-glass of the Oxford University Observatory. We are also indebted to Mr. BELLAMY for the check plates he obtained in January and February.

Thanks are due to the Royal Irish Academy for the loan of the 4-inch object-glass and 8-inch cœlostat.

As stated above, the expeditions were arranged by the Joint Permanent Eclipse Committee with funds allocated by the Government Grant Committee.

[In Plate 1 is given a half-tone reproduction of one of the negatives taken with the 4-inch lens at Sobral. This shows the position of the stars, and, as far as possible in a reproduction of this kind, the character of the images, as there has been no retouching.

A number of photographic prints have been made and applications for these from astronomers, who wish to assure themselves of the quality of the photographs, will be considered and as far as possible acceded to.]

X. *A Critical Study of Spectral Series.—Part V. The Spectra of the Monatomic Gases.*

By W. M. Hicks, *F.R.S.*

Received November 14, 1918,—Read January 23, 1919.

[Plates 2–5.]

Contents.

Page
Introduction 336
Krypton, S series 344
„ links 347
„ double displacements 349
Xenon, S series 350
„ variability with excitation 350
„ links 352
Argon, S series 355
„ links 357
„ examples of displacement 359
D and F series. 360
Krypton, D series 361
„ abnormal D separations 363
„ satellites 365
„ F series 368
„ **F** „ 375
„ value of the oun 377
Xenon, D series 381
„ F and **F** series 383
„ source of 1864, 829 separations 393
„ D series resumed 396
„ D qualification test 397, 406
„ existence of parallel D groups 399
„ satellite sets 402
„ two groups depending on $79\Delta_2$, $70\Delta_2$ 404
„ parallel F and **F** series 411
„ Ritz combinations 421
„ value of the oun 421

VOL. CCXX.—A 580. 3 A [Published May 19, 1920.

Page
Radium emanation 424
„ analogue of 1864 in XF. 426
„ general considerations for S.D. 433
„ value of S(∞) and *e* link 436
„ „ P(∞) and *u.v* links 437
„ S series 438
„ D series 440
„ source of 5648 separations 441
„ value of the oun 442
Argon, D series 443
„ (−5δ) D series 444
„ F and **F** series 445
„ value of the oun 447
„ Analogue of 1864 XF. 449
Neon 449
„ F series 451
„ parallel F sets 452
„ **F** series 453
„ D „ 454
„ S „ 456
„ *e* link 456
„ normal D and S lines 458
„ value of the oun 461
„ WATSON'S separations 463
„ ROSSI'S series 466
Table of constants 468

THE present communication has two objects. Its subject matter is an attempt to obtain some knowledge of the series relations in the spectra of the group of the monatomic rare gases Ne to RaEm, whilst the methods employed will serve to illustrate the fundamental importance as instruments for further research of the new facts brought to light in the previous communications.* The importance of the first object will be generally acknowledged, but it does not yet seem to be realised how definite and exact those new relationships are, even in their as yet undeveloped form, and how powerful an instrument is placed in our hands for the analysis of spectra. It may be well therefore to commence by a brief *résumé* of some of these laws as applied in the succeeding pages. Further, as the establishment of the results obtained must by its nature depend on the numerical comparison of a very large number of lines in all five spectra, and as this evidence must be fully set out to enable a specialist judgment to be formed on it, the communication has unfortunately become very lengthy. The mass of detail will perhaps be rather dreary to the general reader not specially interested in this line of study. It is apt also to hide by its amount and complexity the general conclusions arrived at. I propose therefore to give a slight general survey of these conclusions before giving the evidence.

---

* "A Critical Study of Spectral Series," Part I., 'Phil. Trans.,' A, vol. 210, p. 57.
„ „ „ Part II. „ „ 212, p. 33.
„ „ „ Part III. „ „ 213, p. 323.
„ „ „ Part IV. „ „ 217, p. 361.
These will be referred to respectively as [I.], [II.], [III.], [IV.].

As is well known the wave-numbers of series lines depend on four types of sequences $p(m)$, $s(m)$, $d(m)$, $f(m)$, and that in any one series they depend on the differences between one sequent of one type and the successive terms of the sequence of another type. These sequences are all of the form $N/\{\phi(m)\}^2$, where N is **Rydberg's** constant and $\phi(m)$ is of the form $m+$fraction, the fraction being, as a rule, determinable as a decimal to six significant figures. Our aim is to discover the properties of these functions. The fractional part depends in some way on the order m, although whether it can be considered a definite function of m in the ordinary sense is doubtful.* This fractional part will be referred to as the mantissa, and in dealing with it, it will be regarded as multiplied by 10^6, *i.e.*, as if the decimal point were removed.

The Oun.—It is found† that in each element a constant quantity particular to each element plays a fundamental part in the constitution of the sequences. This is called the oun. The d and f sequences depend in definite ways on multiples of this quantity, whilst it also enters into the constitution of the p and s. Its determination is therefore for each element a matter of the first importance. Denoting its value by δ_1, the quantity $\delta = 4\delta_1$ is of such frequent recurrence that it is useful to treat it as one datum. The oun is accurately proportional to the square of the atomic weight, and is given by $\delta = (361{\cdot}8 \pm {\cdot}1)(w/100)^2$, where w denotes the atomic weight.

In the case of doublet or triplet series, the corresponding separations between them are due to different limits whose mantissæ differ by amounts Δ or Δ_1, Δ_2 (say). In all cases these are found to be integral multiples of the oun. For triplets $\Delta_1 : \Delta_2$ is always somewhat greater than 2.

In the case of D series where satellites occur, the separations of the latter are due to differences in their d sequences. The mantissæ of these latter again differ by quantities which are multiples of the oun, and in the case of triplets they appear in normal types to be very close to the ratio 5 : 3.

The d Sequence.—In the normal type the sequent of the extreme satellite has its mantissa a multiple of Δ_2. The only known exceptions are found in Sr, Cd which show the multiple law, Sr in d_{12} and Cd in d_{11} instead of in d_{13}. In both these cases also the Zeeman pattern is abnormal. As the main lines D_{11} (and in triplets D_{12} also) have their mantissæ greater than that of the outer satellite by multiples of the oun, it follows that all the d sequences for the first order have mantissæ multiples of the oun. It is probable that this is true for all orders of m, but the data are not

* 'Astro. J.,' 44, p. 229, see also [III., p. 339].

† [III.], also 'Proc. R. S.,' A, 91 (1915).

sufficiently accurate to prove this, although they obey the rule within error limits. [III., p. 340.]

The f Sequence.—This sequent of the first order also has the multiple of Δ_2. The material at disposal is not so comprehensive as in the case of the D series, for, except in the second group of the periodic table, the F lines occur chiefly in the ultra-red. The proof of the above statement is perhaps, therefore, not so conclusive as in the case of the *d* sequence. It completely stands the test however in the rare gases. There seems some evidence that F series also show a satellite effect in a small degree —of one or two ouns. In the second group it seems to be a general rule that in many of the low orders ($m = 1, 2 \ldots$) the *f* sequents receive very large displacements from their normal value, so that a normal line is much weaker or is altogether absent and replaced by others separated from it by considerable numbers. This also is found to be the rule in the present case.

Displacement.—Regarding the ordinary doubtlet or triplet series we may consider the second (or third) as displaced from the first by the deduction of a certain number of ouns from the mantissa of the limit; or better perhaps regard the last satellite set as the fundamental one and the others as displaced by the addition of ouns. When such displacements occur in the limit of one line the new one is indicated by writing the displacement on the left. Thus $S_2(m) = (-\Delta_1) S_1(m)$ or $S_1(m) = (\Delta_1) S_2(m) = (\Delta_1 + \Delta_2) S_3(m)$. With satellites, on the other hand, the similar effect is produced in the sequence terms. In this case it is entered on the right. Thus $D_{12}(m) = D_{11}(m)(-x\delta_1)$ or $D_{11}(m) = D_{12}(m)(x\delta_1)$; $D_{22}(m) = (-\Delta_1) D_{12}(m) = (-\Delta_1) D_{11}(m)(-x\delta_1)$. Displacements of both kinds are very common in spark spectra and put themselves specially in evidence in the succeeding pages. A normal line may not only show lines displaced from it, but often it appears to be replaced by them, and, in general, when it does not disappear its intensity is abnormally low. This is practically what happens in the D satellites. The D_{13}, D_{23}, D_{33} appear to be the normal lines in which we should expect descending order of intensity; but most of the energy (or the majority of the emitting centres) appears carried over to the more intense and displaced set D_{12}, D_{22}; and, again, most of what should be expected in D_{12} is carried over to become the strongest line D_{11}. Frequently the D_{13} line has disappeared and only the fragment D_{23}, D_{33} of the triplet left. In general, the D_{11} lines of any element are the strongest of the series. But in the present vacuum tube spectra (spark type) we shall find very frequently that the line required for D_{11} is comparatively weak, and in this case there appear other lines related to it by oun displacements chiefly in the limit. As the real existence of these displaced series is a matter of some importance considerable space has been given in the discussion of the X spectrum (p. 399) to its demonstration in the case of two series depending on the limits $(\pm 2\delta_1) D(\infty)$. It seems a peculiarity of these displacement series that a term of one series may be absent but appear in another. Thus $(-2\delta_1) D(m)$ may not be observed, but a $(+2\delta_1) D(m)$ may be and *vice versâ*. The presence of a similar effect in the sequent terms of F series has been referred to above. One good illustration of double displacements fully established is found in the KrS series (p. 349), in which the indications are shown for $m = 1, 2, 3$. A knowledge of the laws governing displacements is much to be desired. Very little is known at present.

Linkages.—Arc spectra are distinguished, as a rule, by the presence of well-defined series, depending on single groups of P.S.D.F. type. In spark and vacuum tube spectra, however, these seem to be weakened, and a very large number of other lines appear which are related to one another by certain constant separations (links) to form congeries of linkages each connected to a series line. These links can be calculated when the values of Δ, or of Δ_1, Δ_2, are known. The evidence for these was given in [IV.]. There appear to be links of several types. Those already discussed are of two types: (1) separations between successive double displacements of Δ_1 on either side of $S_2(\infty)$ or $p_2(1)$; (2) displacements of Δ_1 on either side of $P(\infty)$ or $s(1)$. Of these, use is confined almost entirely in the present communication to one only of type (1) and both of type (2). They are

$$e = (-2\Delta_1)p_2(1)-(2\Delta_1)p_2(1),$$

or

$$= (-3\Delta_1)S_1(\infty)-(\Delta_1)S_1(\infty),$$

$$u = s(1)-(\Delta_1)s(1), \qquad v = (-\Delta_1)s(1)-s(1).$$

These links themselves may also be subject to small displacements by having their sources on, say, $(x\delta_1)S_1(\infty)$ instead of $S_1(\infty)$. For the present purpose, however, no use can be made of these.

In [IV.] the prevalence of these separations in a spectrum in excess of their occurrence from mere chance was exhibited in a series of curves with abscissæ = separation and ordinates = number of occurrences within a given small amount on each side. Such occurrency curves are also given here for the e links and for the u, v of Kr in Plate 2. The e links seem to be a normal accompaniment to series lines (often displaced, however, when directly attached to those of low order). A further peculiarity of these linkages is the prevalence of the combination $e \pm u$, or $e \pm v$. They are indicated by writing the letter denoting the link to the left of the line when deducted and to the right when added. Thus, in the example below, $44236 = e.47419$, or $47419 = 44236.e$.

Sounding.—In the following pages the unravelling of the complete series of linkages has not been touched upon, but the e, u, v links have been used for testing the existence of lines outside the observed region, a method we may call sounding. A link thus used may be referred to as a sounder. In this way it is possible to obtain evidence of the existence, or of the exact value, of a calculated line which lies beyond the region observed. It may even serve as evidence for the real existence of a line in the observed region too weak to have been observed, for it was shown in [IV.] that the e link appeared to have a tendency to increase the intensity of one of the two lines to which it was attached. The method may be illustrated by an example from KrS. In Kr, $e = 3183\cdot35$. The value of $S_1(5)$ as calculated from the formula is 47419·39, which is in the ultra-violet outside the observed region. But $47419\cdot39 - e = 44236\cdot04$. This is within the observed region, and as a fact the corresponding line is found at 44237·61 with $d\lambda = -\cdot08$ if e is free from error. As an individual case this might be due to a coincidence, but when the same effect occurs with line after line the cumulative effect becomes convincing. To see this it is necessary to get at a glance a survey of all the cases, and for this purpose they are exhibited in sets of diagrams in Plate 3. These diagrams also include links within the observed region in order to show that where the method can be tested it holds. It may be specially noted how the similar arrangement of sounders holds for the same order in the three lines of the same triplet, and how in certain cases the u, v seem alternative. *Cf.* for example XS (1, 3, 4, 7, 8), or the main lines of the three parallel D sets in X, viz. $(2\delta_1)D_{11}$, D_{11}, $(-2\delta_1)D_{11}$, or particularly the prevalence of the $-(e+v)$ combinations in the unobserved lines for KrD. In RaEm these links are too large to be of wide application and in Ne too small to be of use. In RaEm the e link is 23678 and can reach from the unobserved ultra-red across to the unobserved ultra-violet. In Ne the e link is 196, so that its reach is too small to be useful. As this method of sounding is new and clearly of importance if substantiated, considerable attention has been given to its illustration, but as the details themselves are only necessary for a critical study they have been printed in smaller type and may be omitted on a first reading.

Abnormal D *triplet Separations.*—It has generally been held since RYDBERG'S discovery of the satellite systems that the triplet separations for the D and S series are the same. The actual measures did not absolutely prove this, in fact, they indicated small differences, but the accuracy was not sufficient to establish a real difference especially as against a natural bias to expect equality. MEGGERS,* however, has recently placed it beyond doubt that frequently the separations are really different. In Group I., for instance, the separations as measured from D lines are less than those determined from S lines. In the rare gases also this difference appears quite decisively, but here (group 0) the separations as determined from the D lines are, in general, larger than those from S. The key to the explanation is found in the fact that the difference between the two determinations diminishes with increasing order—in other words, that the sequent in the same set of satellites is not the same, and that in a large number of cases the value of $\nu_1+\nu_2$ is the same in both S and D although ν_1, ν_2 themselves are different. It is found to be completely explained by the displacement of one or two ouns between the sequents of the first or second members of a triplet. Sometimes it occurs in the third member. The same explanation accounts for the fact that the F separations are frequently smaller than the corresponding observed satellite separations. It also accounts for the appearance of F satellites as shown here and in [III., pp. 389–395] in connection with the F lines in the alkaline earths. The matter is considered in detail for Kr (p. 363) and is found to hold for the cases of the other rare gases.

---

* 'Bur. Stand.,' Washington, No. 312 (1918).

The Atomic Weight.—It is clear that an accurate knowledge of a first *f* sequent, or of a *d* sequent which belongs to the satellite involving the Δ_2 multiple, gives the means of determining the oun to in general a unit in the sixth significant figure. For these mantissæ are usually of the order of magnitude of 0·8 and are known to six figures. Hence, if the multiple is known, Δ_2 can itself be determined, and since Δ_2 is a known multiple of the oun (determined by the S multiplet separations), the oun is also known to the same degree of accuracy. Further, as the sources of determining it are often quite independent they serve as tests of the determinativeness of the oun itself to the same degree of accuracy. When Δ_2 is considerable, its value is known sufficiently well for it to determine the multiple, and then this exact integer conversely gives the exact value of Δ_2. In the cases of A and Ne, however, the values of Δ_2 are too small to determine uniquely this multiple directly. The difficulty, however, is surmounted by obtaining successively values with increasing accuracy from other considerations until the final test can be applied. As a fact the Ne oun is amongst the most accurate found. Its determination (p. 461) is specially interesting, and indeed is only possible because the material at disposal depends on interferential measures and large accurately known separations. That of X also is a good determination, and is interesting as depending on a number of quite independent data.

As the oun is proportional to the square of the atomic weight within the limits of error of determination of the latter, it is natural to assume that the relation is exact and that $\delta = q.w^2$, where q is a number between 361·8±·1. If this were sustained it would be possible to obtain w with twice the degree of accuracy of the oun and therefore far in advance of any obtainable by chemical methods. In fact the question is raised as to what is actually understood by the atomic weight. Does it refer to the mass of the positive nucleus, or to that and all or a portion of the electrons? The hope might even be entertained of obtaining by this method some knowledge of the number of electrons partaking in the emission of a line if slight changes in the oun could be found. For instance, we shall find in these spectra not a single group of S, D, or F series as in arc spectra, but several independent groups, viz., *d* and *f* sequents, depending on different multiples of Δ_2. If these gave slightly different values of the oun it could be explained by a transference of electrons. There is little evidence of such variation, but it might occur, for instance, in the oun as deduced respectively from Δ_1 and Δ_2. As Δ_1 depends alone on the measurement of the ν_1 separation of a triplet it is not susceptible of such exact determination as Δ_2, and, as a fact, a suspicion sometimes arises that such a slight difference may exist, and that δ from ν_1 is somewhat less than from ν_2 [III., p. 333] as also here.

Suppose the atomic weight is W and the number of electrons involved is xW. Then the oun is given by

$$\delta_x = q\,(\mathrm{W}+x\mathrm{W}/1850)^2 = \delta_0\left(1+\frac{x}{925}\right).$$

If another value depends on y electrons

$$\delta_y = \delta_0\left(1+\frac{y}{925}\right),$$

whence

$$x-y = 925\,\frac{\delta_x-\delta_y}{\delta} = 925\,\frac{d\delta}{\delta},$$

which gives the transference. At present these considerations are only of speculative interest, but a numerical illustration is given below (p. 381) in connection with Kr.

The results obtained in this investigation have given the oun with much greater exactness than any value obtained in [III.], even than that of Ag. The value of $q = \delta/w^2$ has been determined [III., p. 404] as near 361·75 with Ag = 107·88. I now believe from later work that the true value is closer to this than I thought at that time, but in any case it is far less accurate than the ouns themselves. While, therefore, we can use the ouns to give extremely accurate values of the ratios of the atomic weight of the gases, the actual values in terms of Ag are not so exact, although more accurate than those obtained by chemical means. This statement of course depends on the supposition of the exact proportionality of oun and square of atomic weight.

The values of δ as obtained later are here collected and the atomic weight deduced from them by taking q = 361·75.

	Ne.	A.	Kr.	X.	RaEm.
δ . . .	14·4708 ± ·0006;	57·9209 ± ·002;	249·536 ± ·004;	611·0100 ± ·0017;	1787·024 ± ·05
W . . .	20·0005 ± ·0004;	40·0141 ± ·0006;	83·0543 ± 0006;	129·963 ± ·00018;	222·259 ± ·003
Chemical.	20·2	39·88	82·92	130·2	222 to 222·4

It will be seen that in all cases the spectral determinations are much closer to integral values than the chemical, except in the case of RaEm as estimated from HÖNIGSCHMIDT'S value for Ra. In this case, however, the spectral material is defective. It is shown from one of the criteria that a value of the oun = 1785·23 is just possible but improbable, or = 1783·38 almost impossible. These would give respectively w = 222·148 ± and 222·033 ±. It is curious also that from the defective observational work for Ra [III., p. 327] the value of δ from $\nu_1+\nu_2$ = 254096 = 137δ, whence w = 226·43 is also greater than HÖNIGSCHMIDT'S and more in accordance with the value obtained by Mme. CURIE. The value for the Emanation is, however, much more reliable than the above for Ra. If, regarded as a whole, the deviations from the chemical values (RaEm excepted) are greater than chemists will allow possible, it would seem that in this case we are not dealing with precisely the same entity in the two cases.

Special F *Series.*—There appears to be a remarkably stable triplet series of the F type apparent in most of the gases, but more especially evident in X, in which element it was first noticed. Not only are the lines strong and present in a large number of orders, but they appear, at least in X, to be little susceptible to displacements such as are common in other types. The separations are 1864, 829. The occurrency curve for 1864 is shown in Plate 2, fig. 3. In this, in strong contrast to other such curves, it rises to a very high single peak and is practically symmetrical on both sides of the peak. The similar curve for A is shown in Plate 2, fig. 5.

Summation Series.—In the investigation of this XF series a quite new type of series was brought to light. The hitherto recognised series appear as the differences of two terms A − B. The new one has its wave-numbers of the form A + B. In other words, where the old series are difference frequencies the new ones are the corresponding summation frequencies. The notation adopted is to write the corresponding terms in Clarendon type. Thus

$$\mathrm{F}(m) = \mathrm{A} - f(m), \qquad \mathbf{F}(m) = \mathrm{A} + f(m).$$

The list of the lines in X is given on p. 385 up to $m = 30$. For low orders, $m < 3$, the lines are in the ultra-violet and have to be sounded for. Similar summation series coupled with other F series are also common. It probably explains also the crowding of F separations in spectra like that of Cu in short wave regions far beyond the F limit which has always appealed to me as a difficulty. It is possible that summation series may also exist for the P.S.D. series in all elements, but, as a rule, the limits of these are far larger than the F(∞), with the consequence that any **P.S.D.** lines must lie very far in the ultra-violet, a fact which explains why such types if existing have not hitherto been recognised. The existence of these summation series is thoroughly established and their importance as bearing on theories of the origin of spectral lines is evident. They would seem difficult to explain on any of the current theories. But apart from this the existence of the type is of great value for quantitative determinations. This is fully dealt with on p. 384 and it need not be recapitulated here. Its importance for this purpose may be realised when it is seen that it forms the starting point in the analysis of the RaEm spectrum, that it settles in a quite definite way a difficulty arising in the evaluation of the oun in Kr, and that it fixes a very accurate value for the limit of the 1864 series in X, thus simultaneously fixing a particular *d* sequent subject only to observation error in one line.

Groups of D *and* S *Series.*—Not only do we meet with different groups of D series depending on different multiples of Δ_2,* but in the case of Kr there appear to be two sets of lines suitable for S_3—in other words, there are quadruplets. Whilst the two sets S'_3, S_3 give different separations with S_2, and consequently different Δ'_2, Δ_2, they give the same oun, and in connection with them appear two D groups whose outside satellites depend one on a multiple of Δ'_2 and the other on a multiple of Δ_2. It is to be suspected that this is only one example of what may be a common occurrence in spark spectra.

---

* As an example, see p. 403, in X with groups depending on $70\Delta_2$ and $79\Delta_2$.

The order of presentation is generally that in which the investigation was taken. The key was found in obtaining the KrS system, a result first rendered possible by the publication of wave-lengths in the ultra-violet by LEWIS (1915). Amongst them the KrS(1) triplet was found. XS, AS come next in order of definiteness. The spectra of RaEm and Ne are more difficult to deal with, the first because of its defectiveness in range, number of lines, and accuracy, and the latter because of the smallness of its oun and its triplet separations. After the S series of Kr, X, A come the D and F series of Kr, X, the spectrum of RaEm, the D and F for A, and, lastly, the whole sets for Ne. Led by possibly a false analogy to He [I., p. 105], in which doublet series appear in the blue spectrum, the blue spectra were chosen for investigation, and the family group being of even order triplets were looked for. In Ne, with a single spectrum of composite character, the results obtained may have some reference to the red type as well as the blue, especially in connection with certain remarkable constant separations found by WATSON and analogous to the RYDBERG constant separations in the red spectrum of A. One is inclined to think that these red spectra consist mainly of lines of the F type. But the red or first spectra are outside the scope of the present communication. Although it is a very lengthy one as it stands only the beginning of an analysis has been made. The aim has been to lay the foundation for the series framework of this family of elements, and little beyond has been done. The linkages, as a whole, have not been isolated, the red spectra not touched upon, and many interesting effects which will require clearing up are passed over without reference. A great field for investigation is open in these and other spectra for any who are willing to give the necessary time and patience. In some few cases the presentation might have been slightly shortened by merely stating the final result and showing how the necessary conditions are satisfied. But not only would this have passed over certain phenomena of special interest, but one of the objects of the present communication would have been missed, viz., to illustrate the power of the new facts to guide a search even when the details are most bewildering. Moreover, the evidence itself is the more striking when developed from step to step than when the result is directly presented as a finished product.

Krypton.—Krypton shows two spectra, without and with capacity, the former in the red region and the latter further towards the blue. We have measures of some of the stronger lines by RUNGE, and a considerable number of weaker lines, not observed by others, by LIVEING and DEWAR, although the latter are only given to the nearest Ångstrom. The most complete and reliable sets of measurements are by BALY* (red spectrum 6456—3502; blue 5871—2418), and LEWIS† (blue 2416—2145), both of about the same degree of accuracy with probable error in the neighbourhood of ·03 A. Of exact measures there are only two by FABRY and BUISSON‡ for lines at 5870.9172, 5570.2908 I.A. In the red spectrum RUNGE has pointed out constant separations of 945, to which PAULSON§ has added three others. The observations of LEWIS gave me the first clue to the KrS set of lines and thus formed the starting point for the present communication, although a great deal of preliminary discussion of material for this group of elements had been previously done, especially in connection with the separations for certain linkages. In the case of Kr a very large number of separations in the neighbourhood of 786 to 788 had been found, connected also with others of 309, indicating groups of triplets having these values for ν_1, ν_2. Amongst LEWIS' lines a set was found with separations in the reverse order, clearly pointing to a first set of $-S(1)$ or $+P(1)$ lines and corresponding sets for other orders were then easily found. It would seem that there are always a considerable number of separations governed by oun displacements in the limits of the first order, and that of these, three seem to be of a more stable value and correspond to normal triplets. For instance, in all these gases we find a very large number of cases where a S_1 or D_1 line is followed by a line with a separation very close to $\frac{1}{2}\nu_1$. They force themselves on attention on account of their value being so close to the half of a number being sought for, and others may be present although they have not been looked for. In the present case two alternative sets of lines for the S_3 series, one with $\nu_2 = 309$ and the other with $\nu_2 = 341$ appear. In the original search the former was taken because it is reproduced in the D series as well. But later certain difficulties in the determination of the oun, combined with the fact that the corresponding oun multiple in Δ_2, although quite definite, is out of step with the march of their values in the other gases, led me to include the second. This gives a multiple quite in step with the others, and also affords the means of obtaining a good approximation to the oun.

The lines are given in the following table, which also embrace a few obtained by sounding, both wave-lengths and wave-numbers are given :—

---

* 'Trans. Roy. Soc.,' A, vol. 202, p. 183 (1903).

† 'Astro. Journ.,' 43, p. 67 (1915).

‡ 'C.R.,' March 25 (1913).

§ 'Kong. Fys. Sälls. Hand.,' N.F. 25, Nr. 12.

KrS.

	S_1.		S_2.		S'_3.		S_3.
1.	−(10) 2353·95 42468·98	**786·52**	−(10) 2398·38 41682·46	**309·20**	−(3) 2416·31 41373·26	**340·51**	(−1) 2418·13‡ 41341·95
2.	(10) 3778·23 26460·07	**786·45**	(9) 3669·16 27246·62			**341·49**	(4) 3623·74 27588·11
3.	(3) 2489·51 40156·61	**769·80**	(1) 2442·68* 40926·41		‡		‡
4.	(1) 2216·72 45097·86	**774·3**	(1) 2179·3† 45872·2 ± 2	**311·4**	(2) 2164·6 46183·6 ± 2	**786·45 + 339·86**	(2) 2162·7 46224·17 ± 2
5.	(2) 2259·83.*e* (47421·14)					**786·45 + 343·5**	(6) 2299·02.*e.u.* (48551·12)
6.	(4) 2362·18.2*e* ? (48688·12) or (48698·34) §	**788·06**	(1) 2159·5.*e* (49476·18)	**301·54**	(6) 2302·88.2*e* (49777·72)	**786·45 + 340·4**	(8) 2300·35.2*e* (49825·2)

The first three S_1 lines give for the formula

$$n = 51651{\cdot}29-N/\{m+{\cdot}093630-{\cdot}014156/m\}^2.$$

The calculated wave-numbers for $m = 4 \ldots 7$ are in order 45095·25, 47419·39, 48695·38, 49470·47. The first gives O−C, $d\lambda = {\cdot}12$; the others are outside the observed region but are reached by sounding and give O−C values of −·08, ·12, ·22, the errors including those of the sounders.

Quantities relating to the separation 309 will be denoted by dashed letters. With the limit $51651{\cdot}29+\xi$ and separations $786{\cdot}45+d\nu_1$, $309{\cdot}20+d\nu'_2$, $341{\cdot}16+d\nu_2$,‖ the values of Δ_1, Δ_2 are found to be

$$\begin{aligned}
\Delta_1 &= 10969-{\cdot}316\xi+13{\cdot}79\,d\nu_1 = 44\,(249{\cdot}30+{\cdot}314\,d\nu_1-{\cdot}007\xi),\\
\Delta'_2 &= \;\;4244-{\cdot}121\xi+13{\cdot}67\,d\nu'_2 = 17\,(249{\cdot}63+{\cdot}80\,d\nu'_2-{\cdot}007\xi),\\
\Delta_2 &= \;\;4680-{\cdot}121\xi+13{\cdot}65\,d\nu_2 = 18\tfrac{3}{4}\,(249{\cdot}64+{\cdot}73\,d\nu_2-{\cdot}007\xi).
\end{aligned}$$

The value of the oun is thus given by $\delta = 249{\cdot}30+{\cdot}314\,d\nu_2$, $249{\cdot}63+{\cdot}8\,d'\nu_2$, with the uncertainty usually found from triplet separations [see III., p. 332]. From the limit $= p(1)$ and Δ_1 the values of the a, b, c, d, e links are at once found. The result for $e = p_2(1)(-2\Delta_1)-p_2(1)(2\Delta_2)$ is 3183·35. Since $S_1(1) = p(1)-s(1)$,

---

* ? $S_2(3)\,(-9\delta)$, see displaced sets below.

† Is $F_3(2)$, see p. 376.

‡ In region at very end of BALY'S list; 41341 is his last.

§ (6) 2291·26.*e.u.*

‖ Obtained as least square value from $m = 1, 2, 3$, supposing λ equally probable.

$s(1) = 94120{\cdot}27+\xi = \mathrm{N}/\{1{\cdot}079474-5{\cdot}73\xi\}^2$. From this the u, v links, viz., $u = s(1)-s(1)(\Delta_1)$, $v = s(1)(-\Delta_1)-s(1)$ can be calculated. The results are

$$u = 1884{\cdot}03+2{\cdot}33\,d\nu_1-{\cdot}02\xi,$$

$$v = 1942{\cdot}44+2{\cdot}48\,d\nu_1-{\cdot}03\xi,$$

with

$$e = 3183{\cdot}34+4{\cdot}096\,d\nu_1-{\cdot}006\xi.$$

The occurrency curves for these are shown in Plate 2 (fig. 1). It will be seen at once that the maximum occurrences appear with sharp peaks in close agreement with the calculated values.

The calculated lines for $m > 4$ are outside the observation region, but their existence can be indicated by using the above links as sounders. At the same time, in order to give more confidence in the application of the method, the similar linkages are examined for the lines which are observed. It is, of course, understood that the existence of a single link is no conclusive evidence of the existence of the unseen line, as it may be a coincidence. The evidence is cumulative and is seen at a glance in the diagram in Plate 3 (fig. 1). The actual data are given here. In the application of this the frequent link modification must be allowed for, as well as the effect of observation errors. Consequently separations deviating by not more than two or three units from the values determined above are admitted. As a fact, this will exclude a number of real link connections as well as include a certain number of pseudo ones. But by limiting the deviations to these small amounts the conclusions drawn will be more reliable.

SOUNDING Data.

S (1).

(2) 37340			(1) 39799	**1882·57**		
1944·49						
(1) 39284	**3184·49**	S_1	(1*n*) 38497	**3185·23**	S_2	S_3
1882·11						
(3) 41166			(1) 39741	**1940·62**		

Here the u, v links form a series inequality in S_1 with 39284 and a corresponding parallel inequality with S_2. The $-e$ link with S_2 is probably spurious, since 38497 will later be shown to be $D_{15}(2)$ (p. 371), and the link as shown is excessive. It is omitted in the diagram. No links are seen with S_3,

S (2).

(1) 21332						**1884·03**	(1) 29130		(5) 31726
1994·17									**1886·73**
(2) 23276	**3183·57**	S_1	(4) 24062	**3183·70**	S_2	**3183·38**	(1) 30430	**3183·11**	(3) 33613
			1940·68						**1940·93**
(5) 24578	**1881·95**		(2*n*) 26003						(2) 35554

Here is present a chain of e links in S_2 so frequent in Ag and Au [IV.]. Also, as in the case of $m = 1$, are found the same modification effect of about 2 in the u, v links. It may be noted that in addition to the above there are lines (1*n*) 26014·23 and (3*n*) 25295·58 respectively 1232·39 below S_2 and 1232·66 above 24062·92. Now 1231·14 + ·30$d\Delta_2$ is the calculated value of a link *e'* constituted in the same way as *e*, but with Δ_2 in place of Δ_1.

S (3).

(2) 39063 **1879·80**

(2*n*) 33791 **2 × 3182·59** S_1 [S_2] **3180** (1) 44123

(3) 39000 **1942·56**

The line 40926 entered under S_2 is 769·80 ahead of S_1. The *a* link is 768·98, so that it may be $S_1(3)+a$. It is also numerically exactly the displaced line $S_2(3)(-9\delta)$, also $(\delta)S_2(3)$, but see below. Note that there is a rather abnormal $+e$ link with the normal value of S_2. The calculated value of S_3 is 41252·26 and is close on the limit of the region observed by BALY. Again with this, *e'* links are in evidence, viz., (2) 42483·95 and (1) 40023·21 respectively 1231·69 ahead and 1229·05 behind [S_3], whilst the former has $\pm v$ links from it, viz., 1942·92 and −1942·97. They are however not entered on the diagram, since only *e*, *u*, *v* links are there shown.

In the foregoing three orders the links are found in evidence. In the succeeding orders they are used as sounders.

S (4).

(1) 41919 **3178·11** S_1 (1) 42686 **3185·84** $(3\delta_1)S_2$ (3) 43000 **3183·04** $(2\delta_1)S_3$

1943·10

(1) 39976 (1) 43987 **1884·63**

The links here are unsatisfactory as well as the lines given for $S_{2,3}$. Numerically the second and third lines are respectively $(3\delta_1)S_2$ and $(2\delta_1)S_3$. They may really be displacements on the sequence terms, but if so the order is too high to make any certain decision. Taking S_1 as correct, the normal value for S_2 and S_3 would be 45884·31 and 46193·51. The former has links **1881·66** back to (1) 44002·65 with a further *v*, **1942·72**, back to (3) 42059·93. No direct link is found for normal S_3. Its $-e$ link should produce a line at 43010·16. This is very close to the mean of the 43000 shown in the table and a line (4) 43019·17, *i.e.*, 43009·87. In other words, those two lines are $(e)(2\delta_1)S_3$ and $(e)(-2\delta_1)S_3$, indicating that the normal S_3 line has been split into two by displacements $\pm 2\delta_1$ on the limit (or ∓ equal displacements in sequent). Both linked lines are seen, but only one of the lines themselves, which is possibly due to the fact that they are close on the limits of the observed region. We shall find indisputable evidence of such displaced parallel series as a general phenomenon.

S (5).

(1) 42293

1884·65

(2) 44237 **3181·73** [S_1]

1944·48

(3) 42352

The separation $-e$ is given on the value calculated from the formula. In $m = 4$ the observed is 2·6 larger than the calculated, pointing to too small a value for the limit, which is quite possible as the limit determined from lines involving $S(1) = -P(1)$ is never found correctly. An increase of this order makes the $-e$ link in $S_1(5)$ normal. It gives the line in the table with $d\lambda = -·08$.

It is most important to get evidence which can carry conviction as to the reality of displaced lines, especially where the displacements occur simultaneously on both limit and sequent. A numerical coincidence in this case can have little weight by itself. In fact, it can have weight only in each particular case provided it is known from other considerations that such displacements are a common and universal rule. For this purpose it is instructive to adduce here some striking evidence afforded by certain sets of lines connected with the S series. The lines in question are arranged in the two following schemes, expressed in wave-numbers :—

−(1) 42844		−(3) 42059		−(8) 41496	
123·00		**124·53**		**122·92**	
−(2) 42721	**252·37** $S_1(1)$	−(1) 41935	**252·94** $S_2(1)$	−(1) 41625	**252·30** $S'_3(1)$
300·11					
−(1) 42421					

(1*n*) 26407	**35·23**	(1) 26442	(2) 27175			
17·31		**17·29**	**22·53**			
(10) 26424	**35·31**	$S_1(2)$	(1) 27207	**38·70** $S_2(2)$	(1) 27520	**35·17** [$S'_3(2)$]
299·56		**299·72**	**298·15**			
(2) 26724	**35·47**	(5) 26759	(2*n*) 27506			

These lines are numerically displaced with reference to the S by amounts represented by the following parallel schemes :—

$S_1(1)(-9\delta)$		$S_2(1)(-9\delta)$			$S'_3(1)(-9\delta)$
$S_1(1)(-6\delta)$,	$S_1(1)$,	$S_2(1)(-6\delta)$,	$S_2(1)$,	$S'_3(1)(-6\delta)$,	$S'_3(1)$
$(-17\tfrac{1}{4}\delta)S_1(1)(-6\delta)$					

$S_1(2)(-9\delta)$,	$S_1(2)(-3\delta)$,	$S_2(2)(-9\delta)$			
$S_1(2)(-6\delta)$,	$S_1(2)$,	$S_2(2)(-6\delta)$,	$S_2(2)$,	$S'_3(2)(-6\delta)$,	[$S'_3(2)$]
$(-17\tfrac{1}{4}\delta)S_1(2)(-6\delta)$,	$(-17\tfrac{1}{4}\delta)S_1(2)$,	$(-17\tfrac{1}{4}\delta)S_2(2)(-6\delta)$.			

In addition, for $m = 3$, we have seen that in place of normal $S_2(3)$ the displaced $S_2(3)(-9\delta)$ is seen. The parallelism in spite of lacunæ show that the set are definitely related, and the fact that the same displacement on the sequents for $m = 1, 2, 3$ are required to represent the observed separations is specially striking, it being remembered that a displacement on the limit gives constant separation for different orders, whilst one on a sequent gives different for different orders. Here, for instance, 252 in $m = 1$, 35 in $m = 2$, and 16 in $m = 3$, all depend on the same own multiple, 9δ, displacement in the sequences. Also 123 in $m = 1$ and 17 in $m = 2$ on the same 6δ, whilst the constant displacement 300 is explained by $-17\tfrac{1}{4}\delta = -69\delta_1$ on the limit.

Xenon.—Xenon also shows two spectra, without capacity in the red region and with it, extending far into the ultra-violet. Practically the only material at disposal is contained in the extensive lists of BALY* (red spectrum 6198—2536, blue 6097—2414) with an accuracy of about ·03 A. This is supplemented by observations by LIVEING and DEWAR,† especially by longer wave-lengths up to 6596, but unfortunately only measured to the nearest unit.

BALY draws attention to the large number of lines apparently common to both Kr and X. The number of lines in the whole spectrum is very large. BALY gives 1376 in the blue spectrum, but perhaps the most noticeable point for our present purpose is the great variability with change in the conditions of excitation. This is very clearly indicated by a comparison of intensities of corresponding lines as observed by L., D. and B. The following are a few examples out of a large number illustrating this. The numbers following a wave-length give the intensities as estimated respectively by BALY and by LIVEING and DEWAR :—

5191	5, 6	4890	5, 3
5188	4, 3	4887	5, 0
5080	7, 2	4884	1, 4
5068	not seen, 5	4883	6, 0
5045	3, 6	4844	10, 10

As between the two spectra also, a fact noticed by L. and D. is of importance. They say "there is one very remarkable change in the xenon spectrum produced by the introduction of a jar into the circuit. Without the jar the xenon gives two bright green rays at about λ4917 and λ4924, but on putting a jar into the circuit they are replaced by a single, still stronger, line at about 4922. In no other case have we noticed a change so striking." They also state that changes occur with the same kind of discharge as between different tubes. These are clear cases of oun displacements. PAULSON again (*loc. cit.*) gives some constant separations in the first spectrum. The triplet separations observed are about $\nu_1 = 1778$, $\nu_2 = 814$, in due order of magnitude with those for Kr. No line suitable for S (1) comes within the observed region, but there are two lines with W.N. 40375·40, 39561·50 separated by 813·99, which would serve for $S_2(1)$ and $S_3(1)$ and are in a similar position to the KrS lines. They clearly suggest that the $S_1(1)$ line is at −42153·39, using $\nu_1 = 1777{\cdot}90$ as found from the S_2 set. This is further substantiated by employing the value of the *e* link, found below to be 7314, as a sounder. It requires a line at about 34839, and such a line is found at (< 1) 34836·78 (but see discussion under

---

* 'Phil. Trans.,' A, vol. 202, p. 183 (1903).

† 'Roy. Soc. Proc.,' vol. 68, p. 389 (1901); 'Coll. Papers,' p. 494.

S_3(3)). The observed S sets, as well as others found by sounding, are given in the following table :—

XS.

m.	S_1.		S_2.		S_3.
1.	– [2371·57] [42153·39]	**1777·90**	– (10) 2476·02 40375·49	**813·99**	– (4) 2526·97 39561·50
2.	(4) 3854·44 25936·90	**1777·90**	(5) 3607·17 27714·80	**815·11**	(1) 3503·99 28530·91
3.	(4) 2527·10 39559·46	**1776·69**	(1) 2418·47 41336·15	**814·63**	(< 1) 2869·71.*e* (42150·88)
4.	(1*n*) 2689·82.*e* (44480·53)	**1777·44**	(4) 2871·85.*e.u* (46257·97)	**817·68**	(1) 2829·35.*e.v* (47075·65)
5.	(1) 2828·01.*e.u* (46797·58)	**1779·33**	(1) 2944·78.2*e* (48576·91)		
6.	(2) 2452·76.*e* (48072·37)	**1778·33**	(1*n*) 2623·31.*e.v* (49850·70)	**815·34**	(1) 2549·05.*e.u* (50666·04)
7.	(< 1) 2921·74.*e.e* (– δ_1) (48866·68)	**1777·80**	(4) 2777·10.*e.e* (δ_1) (50624·48)	**815·22**	(1) 2715·91.*e.e* (– δ_1) (51439·70)
8.	(2) 2658·37.*e* (– δ_1).*v* (49350·39)	**1778·69**	(1) 2538·16.*e.v* (51129·08)	**816·00**	(2*n*) 2468·54.*e.u* (51945·08)
9.	(1) 2850·41.2*e* (49700·78)				
10.	(1) 2616·79.*e* (– δ_1).*v* (– δ_1) (49949·59)	**1777·55**	(1) 2701·99.2*e* (51727·14)	**815·61**	(1) 2432·87.*e* (– δ_1).*u* (– δ_1) (52542·75)
11.	(1) 2584·04.*e.u* (50134·99)	**1781·25**	(2) 2470·30.*e.u* (51916·24)	**813·57**	(1) 2943·07.2*e.u* (52729·81)
12.	(2) 3202·17.2*e.v* (50276·15)	**1777·00**	(1) 2479·98.*e.v* (52053·14)	**817·59**	(3) 2614·13.2*e* (52870·73)
13.			(2) 2663·43.2*e* (52162·72)	**815·00**	(< 1) 2921·74.2*e.u* (52977·72)

The first three lines gives the formula

$$n = 51025{\cdot}29 - \mathrm{N} \Big/ \left\{ m + {\cdot}096726 - \frac{{\cdot}011826}{m} \right\}^2.$$

From the limit 51025, and using the observed separations given by the S (2) lines, viz., 1777·90, 815·05, the values of Δ_1, Δ_2 are found to be

$$\Delta_1 = 24893 + 13{\cdot}64\, d\nu_1 - {\cdot}72\xi = 40\tfrac{3}{4}\, \{610{\cdot}87 + {\cdot}334\, d\nu_1 - {\cdot}018\xi\},$$

$$\Delta_2 = 10996 + 13{\cdot}33\, d\nu_2 - {\cdot}31\xi = 18\, \{610{\cdot}89 + {\cdot}74\, d\nu_2 - {\cdot}017\xi\}.$$

To a first approximation therefore the oun is given by $\delta = 610{\cdot}88$. The calculated value of the e link from the value of Δ_1 is $e = 7314{\cdot}09 - {\cdot}0056\xi + 4{\cdot}23\, d\nu_1$. The result of examining the lines of the blue spectrum for separations of this magnitude is shown in Plate 2, fig. 2. It is distinguished from those of elements hitherto discussed by showing one definite maximum alone at about 7315·3, although there are indications of the appearance of another peak beyond 7317. The displaced $e(\delta_1)$ link shows a difference of 2·32, so that a second peak might be expected at 7317·62. If the actual value is at 7315·3 it would require $4{\cdot}23\, d\nu_1 = 1{\cdot}2$, or $d\nu_1$ about ·30, $d\lambda = {\cdot}04$ distributed between the two $S_{1,2}(2)$ lines. This is possible with ±·02 on each line, but probably excessive for an error on one of them. Both values are tested as links below for the observed lines, and the results show that with the exception of $S_1(1)$ the value of e, calculated from the original ν_1, is extraordinarily exact. For this reason, and because the exact position of the peak of the frequency curve depends on several disturbing conditions, the original value $e = 7314{\cdot}1$ will be used for sounding purposes on lines outside observed regions. Again the first line −42153 and the limit 51025 gives 93178 as the value of $P(\infty)$ or $s(1)$. From this the values of the u, v links are found. The complete set are

Links.	Changes per δ_1 displacement.
$u = 4133{\cdot}18 - {\cdot}049\xi + 2{\cdot}19\, d\nu_1$	1·77,
$v = 4428{\cdot}00 - {\cdot}061\xi + 2{\cdot}51\, d\nu_1$	1·84,
$e = 7314{\cdot}1 - {\cdot}0056 + 4{\cdot}23\, d\nu_1$	2·32.

The results obtained by sounding are shown at a glance in diagram Plate 3, which embraces orders up to $m = 13$. The cumulative weight of the evidence is overwhelming in support of the general application of this method. The existence of a series parallel to the normal S at a distance $-e$ is proved, whilst the presence of other linked lines is rendered extremely probable by succession of similar linking in the same set, and in neighbouring orders. Compare, for instance, the triplets in $m = 1, 3, 4, 7, 8$ and the sets for $m = 8, 9, 10$.

Detailed Discussion.—In the following discussion the starting point for the consideration of each triplet set is—after $m = 3$—the value of $S_1(m)$ calculated from the series formula obtained above. The sounders are indicated to the left of each observed line and the values O−C in $d\lambda$ are given on the right, the observed or O line being regarded as the observed sounded line+the link as given above. The value entered in the table of S lines above is, however, not the line as calculated from the formula, but the most probable value as deduced from sounding. They are indicated below by asterisks. For the first three orders the values of $d\lambda$ obtained by using $e = 7315{\cdot}3$ are placed to the right of those depending on $e = 7314{\cdot}1$.

S (1).

[−42153·39]				−40375·49				−39561·50			
$-e(<1)$	34836	·15,	·07	$-e(5)$	33061	·01,	−·06	$-e(3)$	32248	−·07,	−·14
$-2e(4)$	27523	·13,	−·02	$-2e(1)$	25747	·00,	−·15	$-e-u(1)$	28113	·04,	−·07
$-e-u(3)$	30700	·37,	·11	$-e-u(1)$	28928	·00,	−·11	$-e+u(3)$	36379	·11,	·05
								$-e-v(1)$	27819	·00,	−·12

In the case of S_1 it appears as if the $e = 7315$ is much superior. But, as it happens, the value of $S_3(3)$ as calculated from $S_2(3)$ by ν_2 is 42151·20, and so is very close to $S_1(1)$. It will be shown that for this the 7314 link gives very close values, and the linkage probably belongs to $S_3(3)$. The $-e-u$ is doubtful.

S (2).

25936		27714		28530	
$-e(3)$ 18622	·00, − ·19	$-v(2)$ 23282	·51	$v(2)$ 32958	·00, ·17
$-e+u(10)$ 22755	·08, ·16	$-v(-2\delta_1)$	·00		
$-2e$[15441·84]					

In S_1 the intensity of 22755 would suggest that its linkage is a coincidence. L.D. give a line 15447 which may possibly be $(-2e)\,S_1$, for their measures are only to the nearest A.U.

S (3).

39559	41336		[42151]	
See $S_3(1)$	$-e(<1)$ 34022	− ·03, − ·10	$-e(<1)$ 34836†	·02, − ·05
	$-e+u(<1)$ 38155	− ·03, − ·06	$-2e(4)$ 27523	·00, − ·17
	$-e-u(1)$ 29892	− ·20, − ·31	$-e-u(3)$ 30700	·22, ·14
	$-e+v(<1)$ 38453	− ·16, − ·22	$-e-u(-2\delta_1)$,,	·01
	$-2e+u(1)$ 30843	− ·14, − ·16		

† $S_1(1)$ and $S_3(3)$ are very close. This sounder probably is correct here and does not hold for $S_1'(1)$, for which it differs by about 3.

S (4).

[44481·06]		[46258·43]		[47073·48]	
$-e(1n)$ 37166	·02 *	$-e(1)$ 38940	·16	$-e(1n)$ 39763	− ·15
				or (<1) 39754	·23
$-e-u(3)$ 33030	·17	$-2e(1)$ 31628	·09	$-e-v(1)$ 35333	− ·08 *
$-e-u(-2\delta_1)$,,	·00	$-e-u(4)$ 34810	·02 *		
		$-e-v(<1)$ 34518	− ·08		

The 38940, 39763 are too far out to be dependent on e links. They could be dependent respectively on $e(-2\delta_1)$ and $e(+2\delta_1)$. Or more probably the lines 38940, 39754 may be $S_2(4)(-2\delta).e$, $S_3(4)(-2\delta).e$.

S (5).

[46799·33]		[48577·23]		[49392·28]
$-e(1n)$ 39491	·28	$-e(1)$ 41257	− ·23	
$-e-u(1)$ 35350	·08 *	$-e-v(<1)$ 36832	·11	
		$-2e(1)$ 33948	·01 *	

39491 is too far out for a link. It would give a reading for $S_1 = 46805·45$. It is curious, however, to notice that we have sounders for a set with this value, viz. :—

46805·45	**1779·80**	48585·01	**815·52**	49400·53
$-e(1n)$ 39491		$-e(3)$ 41271·15		$-e-u(1)$ 37593·25

S (6).

[48072·77]		[49850·67]		[56665·97]	
$-e(2)$ 40758	·02 *	$-e$[42536]		$-e$[43352]	
		$-e-v(1n)$ 38108	·00 *	$-e-u(1)$ 39218	·00 *

S (7).

[48846·60]			[50624·50]			[51439·80]		
$-e$[41532]			$-e$[43310]			$-e$[44125]		
$-2e(<1)$ 34216	·08	*	$-2e(4)$ 35998	−·08	*	$-2e(1)$ 36809	·08	*
$-e-v(2)$ 37105	−·02							

There seems clear evidence of displacement here producing a separation of about 2, which is the same as that by δ_1 on e. The $-2e$ soundings give $S_3 - S_1 = 1777 \cdot 90 + 815 \cdot 12$, which is correct, but 35998 is 2 in the opposite direction. The lines in the list are therefore deduced from this set, using as sounders $e + e(-\delta_1)$, $e + e(\delta_1)$, $e + e(-\delta_1)$. The values of $d\lambda = \cdot 00$ for the set.

S (8).

[49351·72]			[51129·62]			[51944·92]		
$-e$[42037]			$-e$[43815]			$-e$[44630]		
$-e-v(2)$ 37606	·14		$-e-v(1)$ 39386	·02	*	$-e-v(1)$ 40199	·12	
$-e-u(3)$ 37901	·11		$-e-u(3n)$ 39683	−·06		$-e-u(2n)$ 40497	00	*
$-2v(2n)$ 40497	−·11							

The linkage separation gives for the $-e-v$ **1780·83 + 812·41** $= \nu_1 + \nu_2 + \cdot 29$.
„ „ „ „ $-e-u$ **1782·19 + 813·95** $= \nu_1 + \nu_2 + 3 \cdot 19$.

The separations about 1780 are very common and will be discussed more completely under the D series, but here the origin must be a different one. For S (1) the sounder $-e(-\delta_1) - u$ is taken.

The $-2e$ link for S_1 is probably spurious. The line comes under S_3 as well. This is because of the numerical coincidence $\nu_1 + \nu_2 + 2v = 11448 \cdot 95$ and $e + u = 11447 \cdot 28$. It appears in $m = 9, 10$ also.

S (9).

[49699·55]			[51477·45]			[52292·75]		
$-e$[42385]			$-e$[44163]			$-e$[44978]		
$-2e(1)$ 35072	−·05	*	$-2e(1)$ 36842	·24		$-e-u(1)$ 40841	·16	
$-e-v(1)$ 37953	·18							
$-2v(1)$ 40841	·10							

Only one reliable value $-2e$ for S_1.

S (10).

[49949·21]			[51727·11]			[52542·41]		
$-e$[42635]			$-e$[44313]			$-e$[45228]		
$-e-v(1)$ 38203	·16	*	$-2e(1)$ 36998	·00	*	$-2e(2)$ 37916	·17	
$-2v(1)$ 41091	·07					$-e-u(1)$ 41091·56	·13	*

The $-e-v$ in S_1, $-2e$ in S_2, eu in S_3 give $1781 \cdot 53 + 811 \cdot 70 = \nu_1 + \nu_2 + \cdot 28$ the modified ν_1. For S_1 and S_3 the modified $e(-\delta_1)$, $u(-\delta_1)$, $v(-\delta_1)$ are taken. The line 37916 under S_3 is also $(ev) D_{16}(8)$, considered later.

S (11).

[50134·46]			[51912·36]			[52727·66]		
$-e$[42820]			$-e$[44598]			$-e$[45413]		
$-e-u(1)$ 38687	−·02	*	$-e-u(2)$ 40468	−·28		$-2e-u(1)$ 33968	−·08	
$-2e-v(<1n)$ 31076	·05					$-2e-v(3)$ 33674	−·12	

Again the modified $\nu_1 = 1780$ with $eu(\mathrm{S}_1)eu\mathrm{S}_2$. These and $(2e, u)\mathrm{S}_3$ give

$$1781 \cdot 25 + 813 \cdot 57 = \nu_1 + \nu_2 - 1 \cdot 37.$$

S (12).

[50275·68] $-e$[42961] $-2e-v(2)$ 31219 −·02 *	[52053·58] $-e$[44739] $-e-v(1)$ 40311 ·01 *	[52868·88] $-e$[45554] $-2e(3)$ 38242 −·07 *

S (13).

[50385·80]	[52163·70] $-2e(2)$ 37534 ·03	[52979·00] $-2e-u(<1)$ 34216 ·04

These higher orders are necessarily close to high orders of the D series, and many are therefore apparently common to both. *E.g.*, 34216 has been adduced as $(2e)\mathrm{S}_1(7)$ and is also connected with a D line. Also their wave-numbers are now so high that it requires two sounders in series to just reach the limits of the observed region. The later identifications are therefore all doubtful.

Argon.—For the red or non-condensed spark spectrum about 360 lines between 8015 and 2476 together with 16 lines in the ultra-red have been observed. For the blue or condensed spark spectrum the number amounts to about 780 between 6682 and 2050 together with another 40 lines in the extreme ultra-violet between 1886 and 1333. The measures in the red are due to PASCHEN* (ultra-red), KAYSER,† RUNGE and PASCHEN,‡ and EDER and VALENTA,§ and in the blue to KAYSER,† EDER and VALENTA,§ and LYMAN‖ (extreme ultra-violet). In addition we have exact interferential measures in I.A. by MEISSNER¶ for some red lines and measures by BALY** for a few extra lines in the blue spectrum. The red spectrum is noted for the existence of the sets of constant separations discovered by RYDBERG.†† The present communication, however, deals chiefly with the blue spectrum.

The search for the S series in A is more difficult than in the cases of Kr and X. There are an extremely large number of separations of about the same value but clearly distinct. They range round 179 to 188, and, as will be seen later, displacements are very common. The clue is given from the analogous S (1) lines for Kr and X. The only strong triplet lines in the corresponding positions are those given in the following list :—

---

[*] 'Ann. d. Phys.,' vol. 27, p. 537 (1908).

[†] 'Berl. Ber.' (1896), p. 551; 'Astro. Journ.,' vol. 4, p. 1 (1896).

[‡] 'Astro. Journ.,' vol. 8.

[§] 'Denks. Wien. Akad.,' vol. 64, p. 216 (1896).

[‖] 'Astro. Journ.,' vol. 33, p. 107 (1911).

[¶] 'Ann. d. Phys.,' vol. 51, p. 95 (1916).

[**] 'Phil. Trans.,' A, vol. 202, p. 188 (1904).

[††] 'Astro. Journ.,' vol. 6, p. 338 (1897).

AS.

m.					
1.	−(5) 2344·4 42642·10	**179·43**	−(3) 2354·3 42462·67	**75·60**	−(1) 2358·5 42387·07
2.	(5) 3765·463 26549·76	**181·64**	(2) 3739·88 26731·40	**74·74**	(9) 3729·450 26806·14
3.	(1) 2484·1 40244·05	**179·10**	(2) 2473·1 40423·15	**70·39** **75·59**	(1) 2468·8 40493·54 (40498·74)
4.	(2212·7) (45179·30)	**179·50**	(2204·0) (45358·80)	**75·60**	(2200·3) (45434·40)
5.			(2098·5) (47681)		
6.	[2049·5] [48776·74]		[2042·2] [48950·34]		[2038·8] [49031·94]

The observation errors for $m = 1$ and 3 are very considerable, since the measures are only given to ·1 A.U. and ·05 A. produces about $dn = ·8$. Consequently it is possible only to obtain approximate values for ν_1, ν_2. On the other hand, for $m = 2$, where we have very accurate measures, there must be some doubt about the allocation of $S_3(2)$ because its intensity, 9, is so excessive in comparison with the 5, 2 of S_1 and S_2, and the ν_1 separation of 181·64 is greater than observation errors on the lines for $m = 1$, 3 allow. The latter objection, however, can be set aside as it corresponds to the excess ν_1 observed in Kr and X diffuse sets and, as will be found later, in NeS. In these cases $\nu_1+\nu_2$ comes out to be normal. Here, however, the sum is about 1·35 ± too large, and with the S(1) separations the typical $S_3(2)$ would be at 26804·79 or $d\lambda = ·18$, probably of intensity 1, and so overshadowed by the strong line in the list. As will be seen immediately, the linkages will show that this value is preferable. The linkages will also show the probability of a line at 40498 for $S_3(3)$.

The three first S_1 lines give the formula

$$n = 51731·05 - \mathrm{N}\Big/\Big\{m + ·095901 - \frac{·017878}{m}\Big\}^2.$$

For the determination of Δ_1, Δ_2, the oun and the links, reliable values of ν_1 and ν_2 are required. We have seen that the values obtained from the observed sets of lines are subject to large observational errors. Nevertheless that the true value of ν_1 is not far from that shown by S(1) is indicated by the fact that there are several accurate separations of about 179·5, *e.g.*,

(6) 20621	**179·31**	(8) 20800	**179·61**	(4) 20980
(3) 26893	**179·35**	(1) 27072		
(1) 31359	**179·62**	(4) 31538		

of which the first is part of a linkage. With the limit $51731·05+\xi$ and $\nu_1 = 179·50+d\nu_1$, $\nu_2 = 75·60+d\nu_2$, the values of Δ_1, Δ_2 are

$$\Delta_1 = 2519-·073\xi+14\,d\nu_1 = 43\tfrac{3}{4}(57·59+·32\,d\nu_1-·0016\xi),$$

$$\Delta_2 = 1057-·030\xi+14\,d\nu_2 = 18\tfrac{1}{4}(57·91+·77\,d\nu_2-·0016\xi),$$

or

$$\Delta = \Delta_1+\Delta_2 = 62\times 57·70.$$

The oun is thus given by $4\delta_1 = \delta = 57·7$ with some uncertainty owing to inexactness in the observed triplet separations. Its value calculated direct from the atomic weight 39·9 should be $\delta = 361·8\times(·399)^2 = 57·6\pm·14$, the uncertainty being due to the uncertainty $\pm·05$ in the atomic weight. The value of $s(1) = p(1)-S(1) = 94373·10$, from which the u, v links may be calculated. The e, u, v links are found to be

$$e = 719·71+4\,d\nu_1,$$
$$u = 439·47+2·43\,d\nu_1,$$
$$v = 442·67+2·47\,d\nu_1.$$

The examination of the spectrum gives the occurrency curve shown in Plate 2, fig. 4. It shows a very distinct maximum in the region around 720 but little to show the exact position. The values appear somewhat irregular. If, *e.g.*, the ordinate for 719·6 be drawn, it would be the same as for those at 720·5 and 721. It may be noted that if the maximum is taken at 720·4, ν_1 is ·18 larger = 179·68 and the oun calculated from Δ_1 becomes the same as from Δ. But this is rarely the case in triplets. As, moreover, the link 719·7 when applied to the observed S(1, 2, 3) lines gives better agreement than a link 720·4 we shall use it for the purpose of sounding.

The results of sounding are exhibited in diagram form in Plate 3. The details follow:—

S (1).

42642·10		42462·67		42387·07	
$-e$[41922·39]		$-e$[41742·96]		$-e$(3) 41666	$-·06$
$(-7\delta_1)$(2) 41474		$-e-v$(4) 41299	·02	$(-7\delta_1)$(2) 41218	
$-e-u$ 81·86	·06			$-e-v$	$-·03$
$(6\delta_1)$(1) 41488		$-2e$(2) 41023	$-·02$	$(6\delta_1)$(2) 41231	
$e(-6\delta_1)$(1) 43355	·00	$-2e+u$(4) 41460	·11	$-2e$(1) 40949	$-·10$
		$-2e-u$(2) 40585	$-·10$		
		$+e$(3) 43183	$-·06$		

We find in the three sets clear indications of displacements producing separations of 12 or 13. As they appear in different orders such displacements can only arise in the limit. Two cases occur in $m = 1$, viz.,

$S_1(1)-e-u$ and $S_3(1)-e-v$. They land between the lines indicated above, where the separations are 13·76 and 13·59, or the same within error limits. Now $13\delta_1$ on $S(\infty)$ produces 13·26, and the four lines in question are $41474 \cdot 51 = (7\delta_1)S_1(1)-e-u$ or $e.u.(7\delta_1)S_1(1)$, $41488 \cdot 27 = e.u.(-6\delta_1)S_1(1)$, $41218 \cdot 08 = e.v.(7\delta_1)S_3(1)$, $41231 \cdot 67 = e.v.(6\delta_1)S_3(1)$. That is, S_3 repeated from S_1.

S (2).

26549·76		26731·40		26806·14	
$(-3\delta)(4)$ 25817		$(-6\delta_1)(1)$ 27165		$u-5 \cdot 8(1)$ 27239·82	
$-e$ 29·44	·09	u 71·50	−·09		
$(2\delta)(7)$ 25841		$(6\delta_1)(2)$ 27177		$v-6 \cdot 2(1)$ 27242·58	
$2e-v(4)$ 24665	·34	$-u(-7\delta_1)(3)$ 26285	−·05	$-e-12(3)$ 26098·42	
$-e-u(1)$ 25391	−·17	$-(3)$ 26001		$-2e+u(1)$ 25804	·31
		$-e$ 11·98	−·04		
		(3) 22		$-2e+u+v(3)$ 26248	·11
		$-2e(1)$ 25290	·14	$-2e-v(1)$ 24921	
		$-2e-v(3)$ 24845	·58		

Several interesting points emerge from the above.

(1) The links to 26806 as S_3 are all bad. On the other hand, if the links are regarded as correct, the last three point back respectively to 26803·97, 05·34, 03·35, or a mean 26804·22. If S_1 is correct and $\nu_1+\nu_2 = 179 \cdot 50 + 75 \cdot 60$ as assumed S_3 should be 26804·86, the same within the various error limits. This is therefore supporting evidence in favour of the supposition made above that the very strong line (9) 26806 is not itself $S_3(2)$, but that it overshadows the true one, which ought to be a very weak one.

(2) The existence of separations in the neighbourhood of 12 is also very marked. There are lines (4) 26562·08 at 12·22 above S_1, (2) 26744·15 at 12·75 above S_2, and (3) 26098·42 at 12 above the linked line S_3+e. Further, 12, 2 × 12, appear in connection with the different linked systems, showing the existence of sets of lines depending on displacements in $S(\infty)$.

(3) In S_2 the allotted line is 181·64 ahead of S_1 instead of the normal 179·50. The majority of the links from S_1 are to displaced lines, in which δ_1 gives a separation of about 1·03, so that no direct evidence as to normal value of S_2 is directly available, beyond the fact itself that displacement is very prevalent. The two only direct links, $2e$, $2e+v$, would refer back to lines at 26730·34, 27·77, or a mean of 26729·05, which is 179·3 ahead of S_1 and clearly points to the normal S_2. In other words, the normal S_2 is displaced by $(-2\delta_1)S(\infty)$, but the linked lines refer back to the undisplaced normal lines. With the development of our knowledge of the laws of spectral construction, such facts as these may be expected to be of the greatest importance in settling questions of internal constitution of individual vibrating systems.

S (3).

40244		40423		(40498·74)	
$e+6(3)$ 40969		$e+7 \cdot 3(1)$ 41150		$e(2)$ 41218	·02 *
$e+v(3)$ 41407	−·04	$2e+6 \cdot 4-v(3)$ 41426		$e+v(2)$ 41860	·04
$2e(3)$ 41681	·07	$2e(2)$ 41860	·14	$2e(3)$ 41940	−·16
$2e-v(3)$ 41243	·01	$-e-u(1)$ 39265	−·12	(1) 41488	
				$2e-u$ 500	−·10
$-e+12(4)$ 39536	·02	$-v(5)$ 39981	−·08	(1) 41512	
$-2e+u(1)$ 39244	−·01	$-2e+u(3)$ 39420	·16	(2) 40926	
				u 37	·02
				(1) 40949	

Here again the links to the line observed near normal S_3 are bad, whereas the links treated as good refer ack to a line exactly ν_2 ahead of S_2. Further the 12 or 6 separation is again in evidence. Also the links , $-2e+u$, $-e-u$ refer back to lines differing by about 6 from S_2.

S (4).

[45179·30]		[45358]		[45434·40]	
u(1) 45618	−·02	$e-v$(1) 45635	·00	$-e-u$(1) 44996	−·09
$-v$(4) 44735	·07	$-v$(3) 44913	·12	$-e-u$(1) 44275	−·01
$-2e+v$(2) 44175	−·16				

Here the calculated lines agree remarkably with the sounded. The calculated are therefore adopted.

S (5).

[47501·68]	[47681·18]		[47756·78]
	v(1) 48126	−·12	
	$u+6$(1) 48126	·00	

The only links apparent are for S_2, again with the 6 displacement. These lines are all close to the limit of observed region.

S (6).

[48776·74]		[48956·34]		[49031·94]	
$-2e+v$(1) 47784	·16	$-2e-12$(1) 47522	·09	$-e+u$(1) 48753	·06
				$-2e-u$(1) 47155	−·10

The lines are now so far out of the observed region that the links are too small to refer back to well observed regions even if such lines are really existent.

The foregoing discussion of the S series affords evidence of the existence of displacements of about 12 (or 6). This requires further consideration as affording material on which additional knowledge may be obtained regarding the laws which such displacements follow. The presence of the same displacements in successive terms of the series points to a modification on the limit—either a pure displacement, a linkage effect, or, as the separation is small, possibly the difference of two links. The further evidence to follow points decisively to the existence of displacement. Whether they are due to displacements by multiples of the oun on the limit is, of course, not so convincing. The numerical relations are very closely represented on this hypothesis, but in the case of argon the δ_1 is so small that it produces in S_1 separations of 1·03 only. In lines whose wave-numbers lie about 40,000 or greater, this produces changes in λ of ·06 and therefore comparable with observation errors. In the case of S (2) only—wave-numbers of order 26700—does it produce $d\lambda$ = ·15. The measurements here are by KAYSER, whose errors are probably < ·02, almost certainly < ·05, and a close agreement between calculated and observed lines will give evidence of some weight.

$m = 1$. In S(1) cases of the displacement associated with the linked lines have been given above. They also exist in connection with the S lines themselves. Near $S_2(1)$ are 42448·25 and 42473·49, both of intensity 1; the first 14·42 above and the second 10·82 below $S_2(1)$, or a difference of $25·24 = 2 \times 12·62$. Again (3) 42401·45 is 14·38 above $S_3(1)$. Now $3\frac{1}{2}\delta = 14\delta_1$ displacement on $S(\infty)$ produces a separation of 14·42, and of $2\frac{1}{2}\delta = 10\delta_1$ one of 10·30. If, therefore, the observation errors are very small the lines in question are respectively $(-14\delta_1)\,S_2(1)$, $(10\delta_1)\,S_2(1)$, and $(-14\delta_1)\,S_3(1)$. There is no corresponding $(-14\delta_1)\,S_1(1)$, but (1) 42623·93 is $(-18\delta_1)\,S_1(1)$ with $d\lambda = ·00$.

$m = 2$. In S(2) we find direct displacements on the S(2) lines with a whole set of linked lines. They are indicated in the following table :—

(4) 26562·08		**182·07**	(3) 26744·15		**75·60**	(9) 26806·14 (1) 26833·37	19·75
$-e$ (7) 25841	$-·13$		$-e$ (3) 26022			$-e$ (3) 26098	
$-2e$ (5) 25121	$-·09$		e (2) 27464	$-·12$		$-e+24·61$ (3) 26123	
$-3e$ (1) 24399	·30		$-u$ (2) 26305	$-·12$			
$-u$ (3) 26123	·06		$-e-u$ (1) 25582				
$-e+v$ (3) 26285	·00						
$-3e+v$ (3) 24845	·00						

The line 26562 has been already adduced. It is $(-2\delta)\,S_1(2)$ with $d\lambda = -·02$. The lines 26022, 26098 under $(e)\,S_2$ and $(e)\,S_3$ differ by 75·57 or a normal ν_2, and refer back to lines which would give the true triplet separations with the first line 26542.

$m = 3$. From $S_1(3)$ the two lines (1) 40258·62 and (4) 40273·21 are displaced successively by 14·57 and 14·59, the same separation as in the case of $m = 1$. They are $(-14\delta_1)\,S_1(3)$ and $(-28\delta_1)\,S_1(3)$, with O−C given by $d\lambda = ·00$ in both cases. The first has a modified e link $-722·61$ and the second another of 718·46 to observed lines (4) 39536·01 ; (6) 40991·67.

The D *and* F *Series.*

It will be found that the rare gases show a large preponderance of sets of lines which have all the characteristics of belonging to satellites of D series, or to parallel multiplet F series. In other words where, as in other elements, the d sequences show two displacements from the main d_{11} sequence when triplets are in question, the rare gases show a large number of such displacements, both in the d and the f sequences, many of these being due to large multiples of Δ_2. No attempt has been made to determine the whole system of these satellite sets, a problem belonging to an intensive study of individual elements, but a sufficient number have been adduced to prove their existence and to exhibit some of their characteristics.

Krypton.—The following table contains lines allocated to the D series for krypton and discussed in the present communication :—

KrD.

$m = 1$	D_{18}	(1) 19116·44	**788·25** **790·47**	(2) 19904·69 (1) 19906·91	**309·71** **306·95**	(2) 20213·86
		1860·19				
	D_{17}	(1) 19928·46	**786·03** **789·67**	(1) 20714·49 (3*n*) 20718·13	**311·84** **308·20**	(2) 21026·33
		1048·17				
	D_{16}	(2) 20669·36	**788·40**	(5) 21457·76	**308·82**	(3*n*) 21766·58
		307·27				
	D_{15}	[20763·25]	**789·40**	(2) 21552·65	**306·33**	(2) 21858·98
		213·38				
	D_{14}	(2*n*) 20842·89	**788·78**	(1) 21631·67	**307·68**	(4) 21939·35
		133·74				
	D_{13}	(1) 20871·60	**788·75**	(5) 21660·35		
		105·03				
	D_{12}	(<1) 20875·78	**788·28**	(2) 21664·06		
		100·85				
	D_{11}	(6) 20976·63				
$m = 2$	D_{18}	(1) 38037·85	**787·60**	(1) 38825·45	**311·96**	(2) 39137·41 ?
	D_{17}	[38275·82]	**787·44**	(2) 39063·26		?
		[38324·67]	**787·33**	(2) 39112·00	**308·99**	(1) 39420·99
	D_{15}	(1*n*) 38497·23	**787·16**	(1) 39284·39	**306·87**	(2) 39591·26
	D_{13}	(1) 38531·94		?		
	D_{11}	(5) 38560·46				
$m = 3$	D_{15}	(2) 44408·53	**787·99**	(2) 45196·52	**308·07**	(45504·59)
	D_{11}	(3) 44426·67				
$m = 4$		(47077·38) (47080·55)	**786·19**	(47863·57)	**308·15**	(48171·72)
$m = 5$		(48500·71)	**786·88**	(49287·59)	**305**	[49593·55]
$m = 6$		(49345·98) [49349·39]	**787·97**	(50133·95)	**309·89**	(50443·84)
$m = 7$		(49896·60)				

It will be clearer to consider first the D_{11} lines by themselves, and then the satellite lines in each order. From the first three lines the calculated formula is

$$n = 51655{\cdot}56 - \mathrm{N} \Big/ \left\{ m + {\cdot}897262 - \frac{{\cdot}006513}{m} \right\}^2 .$$

The limit is within errors of the value found for $S(\infty)$, but in view of what happens in the cases of X, and RaEm to be considered later, it may be noted that the difference $D_\infty - S_\infty = 4{\cdot}27$ is very close to a $-\delta_1$ displacement in $S_1(\infty)$, which would produce a separation of 4·42. [*Note.*—If $D(\infty) = S(\infty)$, ξ here $= -4{\cdot}27 + \xi$

of S (∞). If D (∞) = $(-\delta_1)$ S (∞), ξ here = $\xi - {}$·15 of S (∞).] The lines calculated from the formula for the succeeding lines are all outside the observed region. For $m = 4 \ldots 7$ they are respectively in wave-numbers

47079·52, 48500·56, 49349·39, 49896·60,

The detailed discussion immediately following will show the evidence for their existence by sounding. The existence of a parallel set at a distance $-(e+v)$ is brought to light which show for the above calculated lines values of O−C = ·09, ·00, ·14, −·01.

The agreement is remarkably close and would seem to show that the value of D (∞) used is very close to its true value, about one or two units in excess. The results of the discussion are exhibited as a whole in Plate 3, fig. 4.

As in the previous cases, the linkings with the observed lines $m = 1, 2, 3$ are given, in order to show that the method is justified where it can be tested.

1.

(6) 20976
e (4) 24158 ·17
v (1) 22914 ·86
$v\,(-\delta_1)$ 22914 ·00
$-e-v$ (3) 15856

2.

(5) 38560
e (1) 41743 ·04
$-v$ (1) 36619 ·27

3.

(3) 44426
$-v$ (10) 42483 ·01
$-2e$ (4) 37061 ·13

4.

[47079·52]
$-e$ (7) 43894 ·08
$-e-v$ (1) 41951 ·09
$-u$ (2) 45196 ·00

5.

[48500·56]
$-e$ [45317]
$-e-v$ (1) 43374 ·00

6.

[49349·39]
$-e$ [46166]
$-e-v$ (2) 44220 ·13

7.

[49896·60]
$-e$ [46713]
$-e-v$ (1) 44771 −·01
$-2e-v$ (3) 41588 ·05

In $m = 1$ negative links lead to lines in the red where observation is defective. LIVEING and DEWAR give a line to the nearest unit which may possibly be a $-e-v$ link. The line 22914 is too far out to be an exact v, but its difference is 4·47, which corresponds to an exact $-\delta_1$ displacement on the limit. In other words, it belongs to the limit as calculated from the S series.

In $m = 4 \ldots 7$ the $-e-v$ sounders all show lines in evidence. The $-e$ sounder for $m = 5, 6$ would show lines just within the boundary of observation, and the absence of corresponding observed lines is therefore explicable. In $m = 4$ the line 41951 is part of an apparent triplet.

(1) 41951·59 **786·19** (1) 42737·78 **308·51** (2) 43045·93.

The diminished values of ν_1, ν_2 indicate that the second and third lines correspond to D_{22} and D_{33} lines. The value of ν_1 is however sufficiently close to make the difference due to observation error, in which case

41951 would correspond to a D_{12}, leaving the calculated 47079 as a true D_{11} ($d\lambda = 0$), whilst D_{12} is 47077·38. That 47079 exists is also shown by the u sounder. In the table this arrangement is adopted. The line entered for $m = 5$ is the calculated, as it is so close to the deduced. For $m = 6, 7$ they are the deduced from the $-e-v$ sounder. The sounders for D_2, D_3, $m = 4...7$, are indicated in Plate 3, fig. 4.

The Abnormal Satellite Separations.—The separations of the lines suggested for the satellite sets show abnormal values in that they are roughly about 2 greater than ν_1 for the S triplets. The difference is real and not due to errors of observation, and we shall find a corresponding abnormality in the other elements of the group. Taking BALY's maximum error to be $d\lambda =$ ·05, the maximum error in n for the D(1) lines will range from ·21 to ·24, or, say, ·45 on a difference of two lines. All the D(1) readings for ν_1 can therefore be the same within observation errors, but cannot possibly agree with that for the S set. Those for ν_2 however, 308·82, 307·68, cannot be the same without allowing errors larger than $d\lambda =$ ·05. If they are to be the same the excessive error is probably in 21766, which is nebulous and would require $d\lambda =$ ·08. Further, in addition to the lines assigned here to the D series, there are a very large number of other lines showing separations of 788. The question arises, therefore, as to the origin of this abnormality, and it is important to discuss the various possible sources. The formula gives so closely the values of the lines for D_{11} from $m = 1$ to $m = 7$ that there can be little doubt as to the essential correctness of the D_{11} allocation. The limit of the series cannot then be very different from $S(\infty)$.

(1) Is 788 a real separation—*i.e.*, is it produced by a displacement on $S_1(\infty)$ by a larger oun multiple—in this case of $44\frac{1}{4}\delta$ in place of 44δ? If so the separation would be 4·52 greater, or 791 instead of 788, and such an explanation is therefore quite inadmissible.

(2) Is it a b link modified by displacement? If $D(\infty)$ be as found, *i.e.*, $(-\delta_1)S(\infty)$, ν_1 will be increased by ·09 or 786·45 to 786·54—an inappreciable change. To produce a change of 2 in the value of b or ν_1 the limit would have to be $(-5\delta)S_1(\infty)$, which gives a value 88·5 above $S(\infty)$. But, as a fact, the limit found is quite close to $S_1(\infty)$. This explanation is therefore excluded.

(3) Are these displacements on the d sequences? In the normal case the d sequences for a given satellite triplet are the same, but are displaced from one satellite set to another. Is it possible that the sequences suffer displacement in the same triplet also? Take, for example, the satellite set whose first line is 20669. The sequent is $d = 51655 - 20669 = 30986{\cdot}20 = N/(1{\cdot}881350)^2$. A displacement of $-\delta_1$ on the denominator 1·881 increases d by 2·06. If the displacement is on the first line of the triplet it must be $-\delta_1$, if on the second $+\delta_1$, and both give practically the same value 788·51 for the apparent separation in general agreement with the observed value. There is nothing, however, to show whether the displacement should be $-\delta_1$ on the d_1 or $+\delta_1$ on both d_2, d_3, as the observed 308·82 is within our assumed error limit, but it is interesting to note that if the oun for ν_2 be the same as for ν_1 the true value of ν_2 would be very close to the observed. In this connection it is to be

remembered that in the d sequences the oun is not affected by the peculiar triplet modification shown by all elements. In the 20842 set the observed ν_2 is 307·68, but the value of the observed $\nu_1+\nu_2$ is very nearly normal. This means that the third line does not suffer displacement, but only the middle one 21631.

If this explanation is correct, the modifications must diminish with increasing order. For instance, in $d(2)$, δ_1 produces a change in separation of ·56, and the new $\nu_1 = 786{\cdot}45+{\cdot}56 = 787{\cdot}01$ as against 787·16 observed. For $m = 3$ δ_1 produces ·23, but the possible observation errors in n—maximum $dn = 1{\cdot}0$—are now so great that the observed separation of 787·99 is well within the limits of 786·68. So far then as merely numerical agreement goes this explanation would seem very satisfactory, but the changes required are so small that by themselves they can give little confidence. We shall, however, see later how it explains certain effects in the $\mathrm{F}(\infty)$ (p. 368)—which depend on the $d(1)$ sequents—and, further, how it also explains similar modifications in the other elements of this group. Meanwhile further evidence in its favour may be obtained from linkage considerations. Some examples follow.

(*Note.*—The observation errors in the separations should not exceed about ·50.)

(*a*) The mesh

	786·03	(1) 20714	**311·84**					
(1) 19928				(2) 21026	**787·36**	(4) 21813	**309·47**	(1) 22123
	789·67	(3*n*) 20718	**308·20**					

Here with 20714 ν_1 is normal, ν_2 abnormal, but $\nu_1+\nu_2 = 1097{\cdot}87$. Our explanation gives $788{\cdot}51+309{\cdot}20 = 1097{\cdot}71$. Thus on the upper set the first and second have the same d sequent, whilst the third has $(\delta_1)\,d$. In the lower set, on the other hand, 20718 has $(\delta_1)\,d$, the same as for the third.

So also in D_{18}. In the first set d has δ_1 in second and third lines, in second set it has $2\delta_1$ in second line.

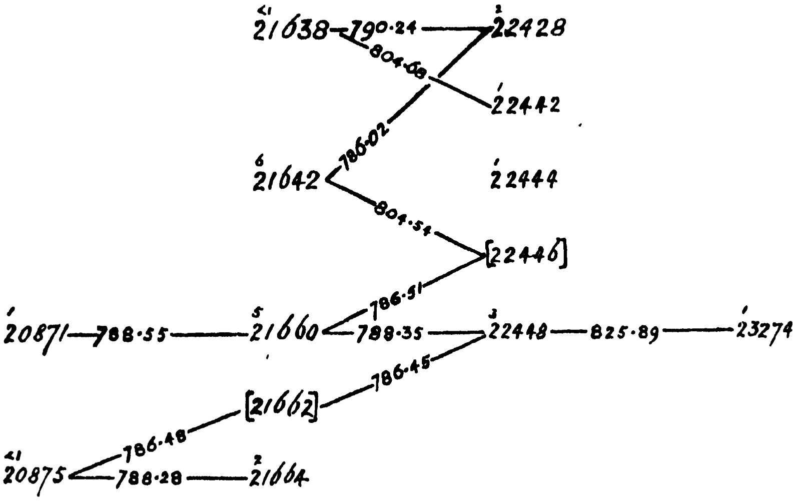
(*b*)

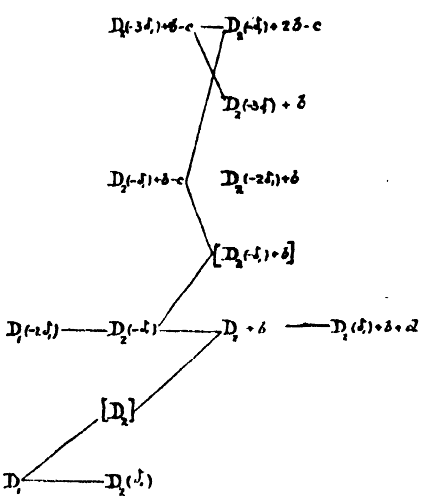
	—**786·43**	(2) 23390	**788·77**—			
				(3) 24178	**308·77**—	
(4) 22603	**789·15**	(4) 23392	**786·03**—			
	—**311·03**	(1) 22914		**2 × 786·47**		(4) 24487

This is a striking example of persistence of the displacement in linked lines. Of the lines in the second column, the first has the same sequence term as 22603, the next two have the links ν_1, ν_2, but the sequent displaced δ_1. The line 24178 keeps the same displaced sequent as these last, and therefore has a normal ν_1 to 23392, but the abnormal to 23390. So 24487 keeps this same δ_1 displaced sequent and so shows a normal ν_2 to 24178 and $2\nu_1$ to 22914. In other words, if the first line is denoted by X, the above sets may be denoted by the following scheme:—

	$\mathrm{X}+\nu_1$		
		$\mathrm{X}(\delta_1)+2\nu_1$	
X	$\mathrm{X}(\delta_1)+\nu_1$		
	$\mathrm{X}(\delta_1)+\nu_2$		$\mathrm{X}(\delta_1)+2\nu_1+\nu_2$

(c) This example involves the c, d links whose values are respectively 804·59, 823·29. It belongs to the D(1) system. The lines enclosed in [ ] are hypothetical and introduced to indicate the transitions. The relations are indicated by the accompanying schematic arrangement, starting with 20875 = D_1.





It may be noted that in this fragment of a linkage all the links involved are p-links. A similar preponderance of these in linkages connected with D lines is a marked feature also in Ag and Au [IV., p. 389].

The conclusion is to be drawn that there is very considerable evidence in favour of (3) as explaining the origin of these modified separations. It does not, of course, follow that the effect indicated in (1) does not exist amongst the lines of a spectrum. We know also that the effect indicated in (2) is existent, and we shall find clear evidence of it in the existence of lines depending on limits which are displaced S(∞), *e.g.*, in the case of 19928 considered below. Further support to (3) is given by the F separations (p. 370).

The Satellites.—Any allocation of satellites which may be regarded as firmly established is a matter of some difficulty on account of the large number of ν_1, ν_2 separations which enter indirectly as links, and the prevalence of sets depending on displaced limits and displaced d sequences. The lines in the list for $m = 1$ are placed there provisionally and for special discussion, although other lines certainly belong to the system. Consequently, the denotation D_{1n} is to be regarded simply as a means of referring to the different sets considered in this communication only, the n giving the ordinal position starting from 20976 as D_{11}.

The criteria that these lines should be possible D satellites with the same D_{11} are that the differences of their mantissæ from that of D_{11} should be multiples of the oun. Should any belong to another group with limit ($y\delta_1$) D(∞), their mantissæ will be modified, but the qualification test will still hold with reference to the difference between their mantissæ thus modified and that of the original D_{11}, for they must all belong to d sequents and so be oun multiples. Nor will the test be affected by the existence of the supposed cause of the abnormal D triplet separations. The value of δ given by both the $S_3(1)$ has been seen to be 249·6 with $\Delta'_2 = 4243{\cdot}2$, $\Delta_2 = 4780{\cdot}0$. The maximum observation errors will be taken as ·03 A and the actual O–C error be written $-{\cdot}03p$. The mantissæ for the various D_1 lines are then

1.	20976	$890749-30{\cdot}82\xi+4p$,	
2.	20875	$887648-30{\cdot}66\xi+4p$,	$3101 - {\cdot}16\xi+4(p_1-p_2) = 12\tfrac{1}{2}\delta-19-\ldots$
3.	20871	$887520-30{\cdot}66\xi+4p$,	$3229 - {\cdot}16\xi+4(p_1-p_3) = 13\delta-16-\ldots$
4.	20842	$886640-30{\cdot}61\xi+4p$,	$4109 - {\cdot}21\xi+4(p_1-p_4) = 16\tfrac{1}{2}\delta-9-\ldots$
5.	[20763]	$884207-30{\cdot}50\xi+4p$,	$6542 - {\cdot}32\xi+4(p_1-p_5) = 26\tfrac{1}{4}\delta-10-\ldots$
6.	20669	$881350-30{\cdot}36\xi+4p$,	$9399 - {\cdot}46\xi+4(p_1-p_6) = 37\tfrac{3}{4}\delta-23-\ldots$
7.	19928	$859253-29{\cdot}30\xi+3p$,	$31496-1{\cdot}52\xi+4(p_1-p_7) = 126\tfrac{1}{4}\delta-16-\ldots$
8.	19116	$835908-28{\cdot}21\xi+3p$,	$54841-2{\cdot}61\xi+4(p_1-p_8) = 219\tfrac{3}{4}\delta-8{\cdot}6-\ldots.$

The test requires that the expressions to the right of the first term in the last column should vanish. The effect of any possible value of ξ is small. Further, as δ is probably known within ±·2, it is only in the two lines, 7, 8, that it can be effective towards satisfying these conditions. It is clear then that the conditions can be satisfied within possible observation errors by Nos. 4, 5, 7, 8, and not by Nos. 2, 3, 6, nor by 8, if δ requires diminution by as much as ·1.

Considering 1, 4, 5, it may be noticed that the separations of 4, 5 from 1 are due to $16\tfrac{1}{2}\delta$ and $26\tfrac{1}{4}\delta$. Since $16\tfrac{1}{2}\times 8 = 132$ and $26\tfrac{1}{4}\times 5 = 131{\cdot}25$, these separations are in the normal triplet satellite separation ratio. Moreover, the mantissa of the extreme satellite $884207 = 189\times 4678{\cdot}3 = 189\Delta_2$ within our present degree of approximation to δ. This again satisfies the normal rule. It is clear, therefore, that these three lines form a normal triplet D (1) set. As to Nos. 7, 8, the ratio of separations is not that of two satellites to 20976 as D_{11}, although near it. The mantissa of (8), however, is $835908 = 197\times 4243{\cdot}1 = 197.\Delta'_2$ within our present approximation to δ, and suggests that it is the extreme satellite of another group. Returning to the other lines it remains to see if they satisfy a test with a displaced D (∞). Now $y\delta_1$ on D (∞) produces a change of $-4{\cdot}42y$ and consequent changes in the mantissæ as follows: $135{\cdot}51y$ in (2), (3), $134{\cdot}18y$ in (6), $124{\cdot}7y$ in (8). As $124{\cdot}80 = 2\delta_1$ it follows that the test for (8) is unaltered and those for 2, 3, 6 become

$$-10{\cdot}71y-19-\ldots = 0\,; \qquad -10{\cdot}71y-16-\ldots = 0\,; \qquad -9{\cdot}38y-16-\ldots = 0\,;$$

whereas in (8) the corresponding condition is not changed, but the mantissa becomes $835924+2y\delta_1 = 197\,(4243{\cdot}1+{\cdot}63y)$. In this case a displacement of one oun in the limit produces the same effect as that of two ouns in the sequent. The test for the others is very closely satisfied by $y = -2$, or the supposition that they belong to a group with a limit $(-2\delta_1)\,D(\infty)$ *or* with a limit $(+4\delta_1)\,D(\infty)$, decreasing by δ the mantissæ of the sequents and requiring

$$\cdot 6 - \ldots = 0\,; \qquad 3\cdot 5 - \ldots = 0\,; \qquad 8\cdot 9 - \ldots\,;$$

which again are easily satisfied within observation errors. With regard, however, to (2), (3), it is not common to find two satellites so close together, and as they both belong to doublet sets we should suspect that they belong to different groups, and in fact the condition is satisfied by regarding the limit of 20875 as displaced by $-\delta_1$ on that of 20871.

We shall consider the conditions more fully later in connection with the more accurate determination of the oun. The preceding results are sufficient to show that (1), (4), (5) are a definite normal D satellite on the basis of $\Delta_2 = 4680 \pm$. That (7) (8) belong to a series probably based on $\Delta'_2 = 4243 \pm 1$, and 2, 3, 6 to a parallel system based on the limit $(-2\delta_1)\,D(\infty)$ or $(\delta)\,D(\infty)$.

I omit details about the satellite *e. u, v* links, but the following points are interesting. The line 19116 affords an example of the two displaced D_2 lines analogous to that given in 3(*a*) as explaining the source of the abnormal ν_1 separations. In addition the linked line $2e + D_{18}$ shows the same effect with two lines (2) 25380·26, (1) 25382·58. D_{16}, D_{36} are curious as showing successive abnormal *u* links. Thus

D_{16}

(1) 18780	**1888·61**	(2) 20669	**1888·37**	(4) 22557

D_{36}

(1) 19881	**1885·31**	(3*n*) 21766	**1885·71**	(3) 23652·29	
					3·69
			1889·40	(3) 23655·98	
					3·31
				(1) 23659·29	

The two lines D_{12}, D_{13} which have been shown to have a relative displacement of $2\delta_1$ show *u, v* links to a mid-line displaced δ_1 from each. Thus (3) 22815·84 is 1942·15 or $v - \cdot 29$ below their mean; whilst (1) 19778·52 is 1883·69 or $u - \cdot 34$ below the mean of D_{22} and D_{23}.

The satellites for $m = 2$ are more doubtful on account of the large changes in the mantissæ produced by observation errors. The line 38531 has a mantissa difference from that of d_{11} of $12\frac{1}{2}\delta$, exactly the same as in D_{13} (1). It should, however, be one of a doublet with the second line stronger. The others show mantissæ differences of 28δ, 103δ, 124δ, $224\frac{3}{4}\delta$, which, allowing for the usual changes in the second order, clearly point to the sets as indicated in the table. There appears, however, no analogue in $m = 1$ for 38324. In this case it is important to notice that the ν_2

separation 308·99 corresponds to the displaced δ_1 sequent for $m = 2$. It thus supports the explanation adopted for the triplet modification.

For reference the multiples of the oun giving the satellite separations are collected in the following table, where d_1, d_2 refer to the first and second line in a satellite set.

$m = 1$.			$m = 2$.		
n.	d_1.	d_2,	n.	d_1.	d_2.
			3	12½	
2	12	12¼	5	28	28¼
3	12½	12¾	?	103	103¼
4	16½	16¾	7	124	
5	26¼	26½	8	224¾	225
6	37¾	38			
7	126¼	126¼		$m = 3$.	
8	219¾	220			
			4	19½	

KrF.—The F lines form parallel series in which the constant separations depend on the satellite separations of the D series. In other words the limits are the $d(1)$ sequents which form the satellites. Now it has been shown above that the abnormal triplet separations which the D series exhibit is probably due to the fact that the d sequent for a given satellite triplet is not the same for each of the three lines, but that they are subject to a displacement of one or more ouns. For instance where ν_1 is 788 in place of 786 (in round numbers) the difference 2 is due to the fact that d_{2n} is not equal to d_{1n} but is $d_{1n}(\delta_1)$. If the strongest line is to be taken as normal, we should expect the d_{3n}, or d_{2n} to be normal rather than d_{1n} as D_{1n} is always weaker than the other lines of a triplet satellite. In this case $F_n(\infty) = d_{1n}(\delta_1)$ and the F separation $= F_n(\infty) - F_1(\infty) = d_{1n}(\delta_1) - d_{11}$ which is less than the observed satellite separation by about 2. As a fact we do find these diminished separations. In searching for F lines therefore we have to examine the spectrum for wave-lengths longer than d_{11} and showing as multiplets with separations the same as the satellite separations or less. In the particular case of Kr these are 100, 105, 133, 213, 307, 1048, 1860,* and we are to expect series which we will denote by F_n with n from 1 to 8. From this point of view it is unfortunate that allowance has to be made for the rule as to the excessive displacements occurring in the lower orders ($m = 2, 3$) and that a complete multiplet, showing all the above separations, is not to be expected. With $D_{11}(1) = 20976{\cdot}63$ and $D_1(\infty) = 51655{\cdot}56 + \xi$, $d_{11} = 30678{\cdot}93 + \xi = F_1(\infty)$. The mantissa of f is in general large, say, beween ·7 and ·99. Consequently the region in which $F_1(2)$ is to be found is where the wave-number is less than $30768 - N/(2{\cdot}99)^2$, or say, <18491, or $\lambda > 5408$. This region is examined for the separations in question and lists made. It is then found that some depend on the same first line, in which case, they clearly refer to F or related lines. Several sets are found connected by one of the ordinary links, which excludes at least one of them as a direct F line. It is now easy to select a few sets from the lists which seem suitable for the F_1 line. This with the given $F_1\infty$ gives $f(2)$, and then RYDBERG's tables give a rough approximation to F (3). It is then only necessary to examine the lines near these for the separations, in order to find the actual F_1 (3). The result of this examination is to show that for the first three orders $m = 2, 3, 4$, the only sets which exist without displacements in any correspond to the separation 301, with F_1 lines (1) 17321·51, ($1n$) 23353·84, (3) 26057·20. The formula calculated from these gives as the limit 30678·64. This is only ·29 less than d_{11} and is thus in very satisfactory agreement with the rule. Using the value of d_{11} as found from the D series with $D(\infty) = 51655{\cdot}56+\xi'$ the actual formula is given by

---

* There may, of course, be others depending on D lines other than those considered in the text.

$$n = 30678{\cdot}93+\xi'-N\Big/\left\{m+{\cdot}877406-577{\cdot}8\xi'-\frac{{\cdot}023916-363{\cdot}3\xi'}{m}\right\}^2.$$

In this if $D(\infty) = S(\infty)$, $\xi' = \xi-4{\cdot}27$, but if $D(\infty) = (-\delta_1)\, S(\infty)$, $\xi' = \xi+{\cdot}15$.

The mantissa of the first line F_1 (2) is

$$865448-107{\cdot}26\xi+16p = 185\,(4678{\cdot}1-{\cdot}580\xi'+{\cdot}034p) = 185\Delta_2{}^*$$

within error limits. The lines have been selected as showing the given separations. Quite independently they give a formula with the proper limit and with the first sequence mantissa a multiple of Δ_2. The evidence therefore for the correctness of the allocation is incontrovertible.

In the consideration of the notation for the various parallel F series it will be necessary to determine what is to be understood by the normal separations. If the latter are to be decided directly from the D_{11} (1) lines, the separations must be those given above. But a glance at the table of F lines will show that there is a considerable variation from these, and indeed from one another—due as we have seen to the great variability in the oun displacements. Especially is this noticeable in the separations given by the D_1 (1) satellites as 213·38 and 307·27. In place of these we find values about 2, 4 or 6 less, corresponding to 1, 2 and 3 oun displacements. The separation 301 is the most frequent displacement from 307·27. Now δ_1 alters the separation by 2·03, and therefore $3\delta_1$ alters 307·27 to 301·18. That these deviations from normal values correspond to real F separations can be seen by their frequent repetition in connection with F lines. See for instance the maps for F, especially F (5), in Plate 4.

---

* $185\Delta_2 = 204\Delta'_2+3\delta_1$.

KrF.

[For brevity F (m) is printed as F in each order.]

$m = 2.$

(1) 17321·51		F_1	(5) 17594·17		$F_1\,(7\Delta'_2)$
(1) 17622·73	**301·22**	$(3\delta_1)\,F_6$	(1) 17695·04	**100·87**	$F_2\,(7\Delta'_2)$
[20504·86]	**e**	$F_1.e$	(5) 20991·16*	**213·64 + e**	$F_5\,(7\Delta'_2).e$
(<1) 20813·91	**309·05 + e**	$(-\delta_1)\,F_6.e$			
			(1) 19134·54		$(-2\delta_1)\,F_1\,(51\Delta'_2)$
(6) 17747·14		$F_1\,(10\Delta_2)$	(1) 19343·33	**208·89**	$F_5\,(51\Delta'_2)$
(2) 17952·16	**205·02**	$(\delta)\,F_5\,(10\Delta_2)$	(1) 19348·50	**213·96**	$(-2\delta_1)\,F_5\,(51\Delta'_2)$
			(5) 20991·16*	**1856·62**	$F_8\,(51\Delta'_2)$
(2) 17972·78		$F_1\,(16\Delta'_2+\Delta_2)\equiv X_1$	[22317·89]	**e**	
(1) 18108·55	**135·77**	$(-\delta_1)\,X_4$	(3) 22448·65	**130·76 + e**	$F_4.e$
(1) 18177·75	**204·97**	$(3\delta_1)\,X_5$	(4) 22531·66	**213·77 + e**	$(-2\delta_1)\,F_5.e$
(2n) 18282·20	**309·42**	$(-\delta_1)\,X_6$	(3) 24178·77	**1860·88 + e**	$(-2\delta_1)\,F_8.e$
[21156·13]	**e**				
(1) 21258·41	**102·28 + e**	$X_{23}.e$			
(1) 22203·47	**1047·34 + e**	$X_7\,(-2\delta_1).e$			
(<1) 23013·63	**1857·50 + e**	$(\delta_1)\,X_8.e$			

$m = 3.$

(1n) 23353·84		F_1	(3) 23418·38		$F_1\,(4\Delta'_2)$
(1) 23659·29	**305·45**	$(d_1)\,F_6$	(4) 23518·73*	**100·35**	$F_2\,(4\Delta'_2)$
(4) 25213·56	**1860·72**	F_8	(5) 25278·06	**1059·68**	$F_8\,(4\Delta'_2)$
(5) 21573·49	**103·68 − u**	$u.F_3$			
(7) 24390·12	**213·00 + d**	$F_5.d$	(2) 23507·07		$F_1\,(10\Delta'_2+\delta)$
(4) 23518·73*	**1048·92 − u**	$u.F_7$	(1) 23639·82	**132·75**	$F_4\,(10\Delta'_2+\delta)$
(2) 23340·79		$(6\delta_1)\,F_1\,(-\delta)$	(1n) 25367·06	**1860·00**	$F_8\,(10\Delta'_2+\delta)$
(1) 23554·41	**213·62**	$(6\delta_1)\,F_5\,(-\delta)$			
(<1) 23349·04		$(2\delta_1)\,F_1$			
(3) 23655·98	**306·94**	$(2\delta_1)\,F_6$			

	$m = 4.$			$m = 5.$	
(3) 26057·20		F_1	[27497·70]		F_1
(1n) 26157·54	**100·34**	F_2	(9) 27708·28	**210·37**	$(2\delta_1)\,F_5$
(3) 26189·73	**132·53**	F_4	(7) 29357·89	**1860·19**	F_8
(4) 26358·25	**301·05**	$(3\delta_1)\,F_6$		———	
(1) 26265·05		F_5-2x	(1) 26727·89		$a.F_1$
[26270·16]	**212·96**	F_5	(7) 26863·38	**135·49**	$a.F_4$
(4) 26275·27		F_5+2x			
(1) 26065·15	**16·05**	F_2+3x	(4) 27482·72	**− 15**	F'_1
	———			———	
(2n) 26067·66		F_1+4x	(4) 27588·11	**105·39**	F'_8
(2n) 26369·36	**304·21**	$(2\delta_1)\,F_6+4x$	(1) 27784·33	**301·61**	$(3\delta_1)\,F'_6$
(1) 27924·30	**1859·15**	F_8+3x	(6) 28535·88	**1053·16**	$(-2\delta_1)\,F'_7$

See also Map F ($_5$).

* At least one a coincidence.

KrF (continued).

$m = 6$.

[28356·47]		[F_1]
(1) 28664·88	308·41	F_6
(2) 26779·99	**307·71 − u**	$u.F_6$
(4) 30218·69	**1862·22**	$(-\delta_1) F_8$
(1) 28361·39	**4**	$(-2\delta_1) F_1$
	———	
(1n) 28574·61	**213·22**	$(-2\delta_1) F_5$
(10) 27525·16	**1047·96 − u**	$u.(-2\delta_1) F_7$
	———	
(1) 28340·49	**−16**	F'_1
(3) 28444·08	**103·59**	F'_3

$m = 7$.

[28907·8] = u.30791·40		[F_1]
(7n) 29005·44	**97·62**	$(\delta_1) F_2$
(2) 30791·40	**u**	$F_1.u$
(3) 30999·10	**207·70 + u**	$F_5.u$
(3) 31841·03	**1049·65 + u**	$F_7.u$
(6) 28891·63	**−16**	F'_1
	———	
(3) 29198·53	**306·90**	F'_6
(1) 30875·06	**99·24 + u**	$F'_2.u$
(1) 31081·96	**306·14 + u**	$F'_6.u$
(6) 31823·11	**1047·29 + u**	$F'_7.u$
(5) 32635·69	**1859·87 + u**	$F'_8.u$

$m = 8$.

[29286·33]		[F_1]
(3) 29498·44	**212·11**	F_5
(7) 28808·81	**309 − b**	$b.F_6$

$m = 9$.

[29552·20]		[F_1]
(2) 29857·70*	**305·50**	F_6

$m = 10$.

[29728·00]		[F_1]	[29712·00]	[F'_1]
(6) 29823·86	**95·86**	$(2\delta_1) F_2$	(3n) 29845·84	F'_4
(1n) 30036·53	**308·53**	F_6	(2n) 29926·23	F'_5
(2) 29857·70†	**129·70**	$(2\delta_1) F_4$	(1n) 30022·82	$(-\delta_1) F'_8$
(1n) 30777·90‡	**1049·94**	$(-\delta_1) F_7$		

* Or $(2\delta_1) F_4$ (10).

† Or F_6 (9).

‡ Is probably $(-\delta_1) F'_1 (7) u$. It is very diffuse and may be both.

It will be most convenient to deal with the multiplets order by order. We are to expect great displacements in the f sequences in the first two orders—single pairs displaced bodily—if the analogy with the triplet systems of the alkaline earths holds in the rare gases. Also throughout the orders we shall find not only in Kr but in the other gases small oun displacements in the limits, owing to the instability of the d sequences.

$m = 2$. In the first order there appear only one normal pair F_1 (2), F_6 (2); but ν_5, ν_6, ν_7, occur in a parallel set linked to the normal by the e link.

The following sets occur amongst others:—

(1)

			1049·06	(2) 18744		
(5) 17594	**100·87**	(1) 17695 ;	**98·37**	(1) 17692·54		
			307·84	(1n) 18002	**97·85**	(2) 18100

Here there is an example of ν_2 or of ν_6, ν_7 appearing as links, omitting 97·85 as a doubtful connection. The ν_2 is the exact D satellite separation. Moreover, the mantissa of 17594 is $895149-110{\cdot}63\xi$, differing from that of $f(2) = 865448-107{\cdot}26\xi$ by $29701-3{\cdot}37\xi+16{\cdot}5p = 7\,(4243{\cdot}00-{\cdot}48\xi+2{\cdot}3p) = 7\Delta'_2$. This is a clear displacement by $7\Delta'_2$, for ξ cannot be more than a few units and p is a fraction. It is clear, therefore, that 17594 is $\mathrm{F}_1(2)(7\Delta'_2)$ and 17695 is $\mathrm{F}_2(2)(7\Delta'_2)$. The next two must therefore be only the linked lines $\mathrm{F}_2(2)(7\Delta'_2)+\nu_6$, $+\nu_7$. There is a line (1) 17692·54, separated by 2·5 from 17695, which might suggest the displacement δ_1 in the limit. But 2·5 is too large, and the line itself is really a linked line to D_{12}, viz., $(e)\,\mathrm{D}_{12}(1)$.

(2)
(2) 17972·78 **135·77** (1) 18108·55; **204·97** (1) 18177·75; **309·42** (2*n*) 18282·20
e
[21154·13] **102·28** (1) 21258·41; **1047·34** (1) 22203·47; **1857·50** (<1) 23013·63

This is a specially interesting displaced set in that it contains all the seven separations, three directly depending on the line 17972 and the remainder on a line linked to it by the *e* link. The mantissa of 17972 is $937966-115{\cdot}6\xi$, which is $72518-8{\cdot}35\xi$ above that of $f(2)$. This is very close to $16\Delta'_2+\Delta_2 = 72553$. The line may be written $\mathrm{F}_1(2)(16\Delta'_2+\Delta_2)$. Call this X_1. Then remembering that δ_1 produces 2·03 in the limit and ·54 in the sequent, we see at once that in the others, for $135{\cdot}77 = 133{\cdot}74+2{\cdot}03$ is $(-\delta_1)\,\mathrm{X}_4$ exact; for $204{\cdot}97 = \nu_5-8$ is $(4\delta_1)\,\mathrm{X}_5$; for $309{\cdot}42 = \nu_6+2$ is $(-\delta_1)\,\mathrm{X}_6$. The others depend on $e+\mathrm{X}_1$. The separation 102·28 is the mean of ν_2, ν_3 which depend on two d sequents differing by $2\delta_1$. It may therefore be written as either $(\delta_1)\,d_{13}$ or $(-\delta_1)\,d_{12}$, we will write it $\mathrm{X}_{23}+e$. For $1047{\cdot}34 = \nu_7-1$ the displacement is $-2\delta_1$ in the sequent, or the line is $\mathrm{X}_7(-2\delta_1).e$. For $1857{\cdot}50 = \nu_8-2{\cdot}6$ and the line is $(\delta_1)\,\mathrm{X}_8.e$.

(3)

$$\begin{array}{rll}
 & & (1)\ 19343{\cdot}33 \\
\begin{array}{r} \\ (1)\ 19134{\cdot}54 \\ \mathbf{822{\cdot}11}\end{array} & \left\{\begin{array}{l}\mathbf{213{\cdot}96}\\ \\ \mathbf{1856{\cdot}62}\end{array}\right. & \begin{array}{l}(1)\ 19348{\cdot}50\\ \\ (5)\ 20991{\cdot}16\end{array} \\
(1)\ 19956{\cdot}65 & \mathbf{308{\cdot}13} & (1)\ 20264{\cdot}78
\end{array}$$

With $30678+\xi$ the mantissa of 19134 is $216805-133{\cdot}50\xi$ above that of $f(2)$. It is in the neighbourhood of $51\Delta'_2$. The whole set of lines are representable as follows (putting as before $\xi = -1{\cdot}34$):—

$$\begin{aligned}
19134 &= (-2\delta_1)\,\mathrm{F}_1(2)(51\Delta'_2-2\delta_1) &&= (-2\delta_1)\,\mathrm{Y}_1(-2\delta_1)\\
19348 &= (-2\delta_1)\,\mathrm{F}_5(2)(51\Delta'_2) &&= (-2\delta_1)\,\mathrm{Y}_5\\
19343 &= \mathrm{F}_5(2)(51\Delta'_2) &&= \mathrm{Y}_5\\
20991 &= \mathrm{F}_8(2)(51\Delta'_2) &&= \mathrm{Y}_8\\
19956 &= (-2\delta_1)\,\mathrm{F}_1(2)(51\Delta'_2-2\delta_1)+d &&= (-2\delta_1)\,\mathrm{Y}_1(-2\delta_1).d\\
20264 &= (-2\delta_1)\,\mathrm{F}_6(2)(51\Delta'_2-2\delta_1)+d &&= (-2\delta_1)\,\mathrm{Y}_6(-2\delta_1).d
\end{aligned}$$

The numerical agreement is very close, and 19956 is a series inequality in which a change of 1 makes the two separations 822, 308 exact d and ν_6 links. On the other hand, as we have already pointed out, 20991 may be a D_{11} line depending on the limit $(7\delta_1) D(\infty)$.

The following points should be noticed in the foregoing allocations :—

(1) The presence of the large displacements in the sequence term by quantities differing from multiples of Δ_2 by one or two ouns—in this respect quite analogous to a corresponding effect in the alkaline earths. It seems to point to a kind of satellite effect in the F series analogous to that shown in the D_1, where the main strong line is displaced from the normal satellite depending on a multiple of Δ_2 by the addition of a few ouns. In this case the F satellite is in general too weak to be observed, except possibly the linked line $22203 = X_7(-2\delta_1).e$ depending on $f(2)(17\Delta'_2)$.

(2) That where a multiplet line is absent, it frequently appears as a linked e line, but that the linked line never appears directly linked to the F_1.

(3) The line 20991 occurs twice as $F_5(7\Delta'_2).e$ and as $F_8(51\Delta'_2)$, also its possible existence as a kind of independent D_{11} line has already been referred to. It cannot of course be all, and at least two of the suppositions must be due to chance. It is also probable that such coincidences may occur in some of the other allocations. The evidence for the general effect is cumulative and not dependent on a single numerical agreement.

The F system of the first order ($m = 2$) have been considered in rather considerable detail in order to establish what appears to be a very general rule that in many groups of elements the configurations producing the normal F lines appear to have been subjected to a sort of explosive effect whereby other configurations producing f sequents displaced by large multiples of Δ_2 are produced. As a natural result the intensities of the normal lines in the spectrum are diminished since the observed intensities must depend on the number of emitting centres as well as the energy emitted by each. We have seen that they are displaced in pairs or sets containing one displaced F_1 line, but no attempt has been made to search for sets not containing the F_1. As we shall see later these displaced sets in the lowest order give a means of obtaining very accurate data for the determination of the value of the oun. In dealing with subsequent orders such a detailed discussion is not called for for this purpose.

$m = 3$. The normal lines are observed for F_1, F_6, F_8, and as illustrating the correctness of the explanation given above for the diminished separation 301 in place of 307 it will be noticed that the normal D value is shown by the line 23349. The other F lines of the set do not directly appear, but as in the case of $m = 2$, they are in evidence as linked lines. Some of the lines linked to $F_1(3)$, are represented in a map in Plate 5. The denotations of the lines are entered in place of the wave-numbers, which can be reproduced by adding the given separations, and each can be referred to by the column and order in the column in which it occurs. Again we have several examples where small changes (errors or displacements) in lines make all the allied links take their practically exact values, and so give evidence for the reality of each. Thus, $-1\cdot2$ on **a1** give c, ν_8; $-1\cdot5$ on **a2** give u, ν_7 and the mean of ν_2, ν_3. (or the line $(-\delta_1)\,F_2$ or $(\delta_1)\,F_3 = F_{23}$, say); $-\cdot5$ ($d\lambda = \cdot08$) on **c4*** gives c, ν_8; $-1\cdot5$ on **c5** gives d, ν_4, ν_8. These new sets give us representatives of all the lines missing from the direct normal lines. It is noticeable how the f_8 sequent persists.

Certain displaced sets are given in the tables. If 23418 is $\cdot5$ less ($d\lambda = \cdot08$) the ν_2, ν_8 become exact and it is $F_1\,(3)\,(4\Delta'_2)$. The numerical proofs of these allocations are not given, as these displacements have no importance at present beyond the fact that they exist.

$m = 4$. Direct lines are found for F_n, $n = 1, 2, 4, 6$, whilst 5 appears displaced $\pm5\cdot01 = 2\times2\cdot50$, to observed lines 26265, 26275. There is also a line 26067 ahead of 26065 by $2\cdot51$. If this $2\cdot50$ be due to some displacement it is probably $2\delta_1$ on the limit and some ouns on the sequence, or all by ouns on the sequence. The order is so high ($m = 4$) that it is not possible to decide, and it is shown in the tables as a difference $x = 2\cdot50$. Linked lines are shown in the map (Plate 5). Again note that $2\cdot5$ on F_1 makes the v, u links exact, and that here again the x appears. Symmetry would seem to indicate that the true $F_1\,(4)$ or 26057 should be about x less. This would diminish the calculated limit of the series to a value nearer that given by the calculated $S\,(\infty)$.

For $m = 5 \ldots 10$ the values for F_1 calculated from the formula are 27498·81, 28357·47, 28909·97, 29286·33, 29554·21, 29730·00. With $\xi = -1\cdot34$ as determined later, we should expect values less than these from about -1 for the first varying to -2 for the last. None of these appear but they have linked lines whilst other of the parallel F sets also appear directly.

$m = 5$. No line has been observed at 27498, but there are lines with it for F_5, F_8, and others for F_{137} by a link $-a$. A value of F_1 27497 is $F_8 - 1860\cdot19$ and reduces it $1\cdot1$ as just suggested.† The connections are exhibited in the map (Plate 4). From this order and beyond there appears to be a parallel set at a distance 16 units less. For $m = 5$, this starts from $27482\cdot72$ as F_1. As is seen in the map (**c9**) it has a very large linkage to F lines with similar sets to those connected with the calculated F_1. We may explain its source as a displaced $(2\delta)\,F\,(\infty)$, as the difference of two p-links, $b-c$, or as the direct congeries of F lines depending on 20991 as an independent D_{11} line. In the map the notation depending on the second is adopted. In the list 27482 is written as F'_1 leaving the question of the origin open. The p-links are particularly prevalent. This was found to be the case also in Ag and Au, the only elements in which the linkages have been examined with any thoroughness. In particular the series of successive + and − links from 27497 recalls a similar arrangement in the AgD (4) linkage shown in the *c*, *d*, *e* columns of the map for AgPiii [IV.]. The series is in fact continued further than is shown in the present map. Starting from 26727 we find $a-c+b-d+a-c+a-c$ (and $+d$) $= 3a+b-(3c+d)$, the actual separations being 770·92 − 803·607 + 787·51 − 821·81 + 771·07 − 803·66 + 770·68 − 803·16 (and + 825·35). Further, it should be noted that each successive pair is a parallel inequality, one in excess and the other in deficit of normal value. It means an increased displacement $2\delta_1$ in each alternate line. But if the observation errors are small, there appear to be indications of simultaneous displacements in the *f* sequences as well as on the limit. In fact a similar phenomenon is indicated in the two next orders though naturally some elements are wanting. A precisely similar connection is shown by AgS (3), [IV., p. 382], in a still more striking and regular series of changes. The elucidation of the laws governing displacements is of the first importance and should be one of the immediate objects of investigation. For this purpose examples of continuous series of simultaneous and like displacements will be of the utmost value. For this reason maps of certain near lines (Plates 4, 5), are given for all the orders from 3 up to $m = 8$, but no attempt has been made to indicate exact displacements involving unity. The parallel series F′ about 16 below F exist for $m = 5, 6, 7$. The sets connected with F′ (7) all show the displacement unity. In the lists the true lines are entered as 1 less for $m = 6$ and 2 less for 7, 8, 9, 10 (*i.e.*, ξ about −2) than the calculated values. As is seen it makes the observed separations more normal and in so far supports the putting of the limit about 2 less. Later the actual change in the limit is found to be −1·34.

---

* This has been given as a bad $D_{11}.e$. The suggested change makes the e link worse, which increases the improbability of its belonging to the D system.

† The calculated is retained however in the map, as the links show the repetitions more clearly.

Kr**F**. During the work of examining the X spectrum a new type of series, associated with the F series, came to light. Whilst the known F type depends on the differences of two sequences $d(1)-f(m)$, the new type has a series of lines whose frequencies are given by $d(1)+f(m)$. We shall denote the lines of these series by **F**, so that F will denote a difference frequency and **F** a summation.

We have already referred to the general properties of these series in the introduction. Some of the material from the Kr spectrum bearing on the subject are here collected. In the following lists each order is considered by itself. The examination has not been exhaustive so as to involve displaced values, but it is believed all the direct observed lines have been included. A few abnormal ones, with considerable displacement in the *f* sequence, have also been entered, as they raised questions which require future investigation. The F and **F** lines are arranged in parallel columns. The mean of the two corresponding lines is entered in thick type between them. That for the first corresponds to the fundamental limit. The succeeding ones are given in the form mean of the first + difference, and the difference only (which settles the denomination of the set) is entered. Thus for $m = 2$ the first mean is 30674·77, that for F_6, $\mathbf{F}_6$ is 30976·55 = 30674·77 + 301·78 and 301·78 is entered. Also over each line the difference from F_1 or $\mathbf{F}_1$ is entered. Notes on detail are appended below the lists. The evidence is clear as to the existence of a series of the form $A+f(m)$. If the limit were the same for both we should expect the mean to be 30678·93 + ξ instead of 30674·77. The latter is what should be expected if D (∞) = S (∞).

TABLE of F and **F** Lines.

	F.		**F.**	F.		**F.**
		m = 2.			m = 5.	
1.	17321	**30674·77**	(2) 44028·04 = (3δ_1)**F?**	[27498]	**30678·23**	(4) 33857·67
	53		**148·79**			
2. 3.	17376	+ **102·10**	(1) 44176·83			
			133·77			
4.			(2) 44161·81			
	213·27		**209·68**			**214·23**
5.	*e*.20718	+ **211·58**	(8) 43433·13 *c*			(1) 34071·90
	301·22		**302·34**			
6.	17622	+ **301·78**	(3) 44330·38			(2) 33838·09 = $\mathbf{F}'_1$
	1033·85		**1052·14**			
7.	18355	+ **1043·00**	(1) 45080·18			(1) 34146·82 = $\mathbf{F}'_5$
	1872·65		**1844**			
8.	19194	+ **1858·40**	(1) 45872·18			
(for the displaced sets see p. 380)						
		m = 3.			m = 6.	
1.	23353	**30677·27**	(2) 38000·71	[28357·47]	**30675·60**	[32993·80]
		10Δ'_2 + δ.				
1.	23507	**30674·76**	[37842·45]			
2.						
			104·03			
3.			(1) 37946·48			
	132·75					**133·70**
4.	23639					(2n) 33127·50
5.						
			306·41	**307·41**		**303·46**
6.			(4) 38148·86	28664		(1) 33297·26
			1048·54			
7.			(1) 38890·99			
	1860·00		**1859·62**	**1860·22**		
8.	25367	+ **1859·81**	(1) 39702·07	30218		
		m = 4 (F′).			m = 7.	
1.	26067	**30677·16**	(6) 35286·68	28907	**30674·71**	32441·62 = *e*.35624·97
						33442·54 = *b*.33228·89
2.						
				F′.		
3.				28891·63		
			129·78			
4.			(1) 35416·16			
			207·17			
5.			(6) 35493·85			
	301·70					
6.	26369					
			1043·80			**1048·78**
7.			(1n) 36330·48	[F′(7)]	**30677 34**	(1) 33511·28
	1856·6		**1859·18**			
8.	27924	+ **1857·92**	(1) 37145·86			

$m = 2$. To the only two direct F there appear two direct $\mathbf{F}$ with the same mean 30674·77 and modified separation 301·8. Also there is a direct line to $\mathbf{F}_4$. To the linked $F_4.e$ corresponds a linked $c.\mathbf{F}_4$. In this connection it must be remembered that all the $\mathbf{F}$ are large, over 44000, and an e link would reach to lines outside the region of observation. Three other sets are included in the list, which involve displaced f sequents. The second pair give a mean 30674·77 + 102·10 and belong therefore to the limit midway between F_2 and F_3. But the sequent = half the difference = 13399·95 in place of 13353·26, on the supposition of a common limit. So also the pairs for $F_{7,8}$ show f sequents 13362·41 and 13339·01 on the same supposition.

If the observed 44028 is really $(3\delta_1)\mathbf{F}_1$, where $\mathbf{F}_1$ has the same limit as F_1, and may be called the normal F_1, the mean of the observed F_1 and of $\mathbf{F}_1$ is 30677·80, and should give the true value of the limit subject only to observation errors on the two lines, *i.e.*, within maximum error of 1·1 with $d\lambda = \pm{\cdot}05$, and within a probable error much less.

$m = 3$. $\mathbf{F}_1$ corresponds to F_1 with mean 30677·27 ± 1. There are some cases of displaced $f(3)$ as in $m = 2$. The $10\Delta'_2$ set appear also in $\mathbf{F}$, and as they contain several examples they are placed in the list. There are two lines 37836·36 and 38050·15 (separation 213·79) which as F_1 and $\mathbf{F}_5$ give a mean 30671·71. The lines in the list show an unobserved line for $\mathbf{F}_1$, which is the basis for the others, its actual value is taken as $-3\delta_1$ displacement on 37836. The mean is 30674·76 which, on the supposition adopted above, corresponds to $a(3\delta_1)\mathbf{F}(\infty)$. Since $F_8 = (-2\delta_1)F_2$, the line 37946 may be written as normal $\mathbf{F}_2$, giving mean limit = 30676·79 ± 1. In the $4\Delta'_2$ set is a line 39799·89 = $\mathbf{F}_8(4\Delta'_2)$, giving with $F_8(4\Delta'_2)$ a mean 1860·19 + 30678·78.

$m = 4$. There appear no direct $\mathbf{F}$ to the F lines. But they occur in the parallel set F′; but as $\mathbf{F}'_1$, $(2\delta_1)\mathbf{F}'_4$, $(3\delta_1)\mathbf{F}'_5$, $(2\delta_1)\mathbf{F}'_7$, $\mathbf{F}'_8$. The mean is 30677·16.

$m = 5$. Here $\mathbf{F}_1$ and $\mathbf{F}_5$ appear, but as the mean depends only on calculated F_1 it is not reliable. If F_1 be taken from the observed line 26727·89 by the $-a$ link, the F_1 line would be 27496·83, giving mean 30677·25. There are also lines connected with the parallel series F'_1 which has a limit 16 below F_1. F'_1 = 27482·72 and $\mathbf{F}'_1$ = 33838·09, gives mean 30660·40, which is about the proper amount below $F(\infty)$ = 30677·80. With this goes 34146·82 as $\mathbf{F}'_5$ with separation 308·73.

$m = 6$. The unobserved lines supposed for the first pair are calculated respectively from the observed F_6, F_8, and $\mathbf{F}_4$. The line 33297 is $(2\delta_1)\mathbf{F}_6$. Corresponding to 33076·55 as $\mathbf{F}'_2$, the mean limit with F'_1 is 30658 in place of 30660. With this might possibly go 34018·93 as $(3\delta_1)\mathbf{F}'_7$.

$m = 7$. $\mathbf{F}_1 + e$ = 35624·97 gives a mean limit 30674·71. Also with $\mathbf{F}_1 + b$ = 33228·89 gives a mean limit 30675·12. Also 33511·22 an exact $\mathbf{F}'_7$ with mean 30677·34.

$m = 8$. $\mathbf{F}_1 + e$ = 35249·74 gives mean 30676·36, but $\mathbf{F}_1$ is uncertain.

$m = 9$. I have not found $\mathbf{F}_1$, but 32015·37 as $\mathbf{F}_5$ gives $\mathbf{F}_1$ = 31802·00, which gives mean limit 30677·10.

$m = 10$. No $\mathbf{F}_1$ found, but 31726·38 as $\mathbf{F}_2$ and 31841·05 as $(-\delta_1)\mathbf{F}_5$ gives $\mathbf{F}_1$ the same value 31625·5. This gives a mean limit = 30676·75.

The evidence seems therefore clear for the existence of this type of series.

The Value of the Oun.—For the evaluation of the oun there are at disposal:—

(1) The Δ_1, Δ_2 as determined from the S separations. These have given (p. 346) for a first approximation to δ, the value 249·30 from ν_1 and 249·6 from the two alternative ν'_2, ν_2. The ν's are so ill-determined that these might possibly refer to values giving the same δ. But the fact that the value of e calculated from Δ_1 agrees so closely with the maximum ordinate in the corresponding occurrency curve (Plate 2, fig. 1) shows that Δ_1 must be exceedingly close to the true value, in which case it is

improbable that the ·3 difference in δ could be attributed to a single observation error on each of Δ'_2, Δ_2. As it has been shown in [III., p. 332] that the triplet separation always shows a slight difference in the δ from ν_1, ν_2, it is probable that the same occurs here also. The evidence there given goes to show that the value obtained from $\Delta_1+\Delta_2$ is always closer to the true value. We should expect, therefore, a value between the two values above.

(2) The evidence obtained from the D qualification test.

(3) The D satellites whose mantissæ depend on multiples of Δ_2, viz., 19116 on $197\Delta'_2$, and 20763 on $189\Delta_2$.

(4) The mantissa of $f(2) = 185\Delta_2$.

Before however conditions (3), (4) can be applied it is necessary to obtain if possible a closer approximation to δ from (2). The material for discussion is that given on (p. 366). We shall discuss it on the two bases of $\delta = 249{\cdot}60+x$ and $249{\cdot}30+x$ where x is certainly not greater than ·3. The complete conditions are, using the displaced values $(-2\delta_1)$ D (∞) for (2), (6) and omitting (3) as parallel to (2),

$$
\begin{array}{lrr}
 & 249{\cdot}60+x & 249{\cdot}30+x \\
(2) & -2{\cdot}4+{\cdot}16\xi-4(p_1-p_2)+13\tfrac{1}{2}x=0 & -6{\cdot}5+\ldots\qquad=0, \\
(4) & 9{\cdot}0+{\cdot}21\xi-4(p_1-p_4)+16\tfrac{1}{2}x=0 & 4+\ldots\qquad=0, \\
(5) & 10{\cdot}0+{\cdot}32\xi-4(p_1-p_5)+26\tfrac{1}{4}x=0 & 2{\cdot}1+\ldots\qquad=0, \\
(6) & -2{\cdot}7+{\cdot}46\xi-4(p_1-p_6)+38\tfrac{3}{4}x=0 & -14{\cdot}3+\ldots\qquad=0, \\
(7) & 16+1{\cdot}52\xi-4(p_1-p_7)+126\tfrac{1}{4}x=0 & -21{\cdot}8+\ldots\qquad=0, \\
(8) & 8{\cdot}6+2{\cdot}61\xi-4(p_1-p_8)+219\tfrac{3}{4}x=0 & -56+\ldots+220x=0.
\end{array}
$$

It is quite clear that the conditions in the first column cannot be satisfied without assuming very large observation errors unless x is negative, nor on the right hand column unless x is positive. In other words, δ must be $< 249{\cdot}60$ and $> 249{\cdot}30$. The first four equations, however, give no indications of amount, as the multiples of x are not sufficient to make the term in x more important than the error terms. In (6, 7, 8) the conditions may be written with $\xi \ngtr 1$.

$$
\begin{array}{lrl}
(6) & -2{\cdot}7\pm8{\cdot}5+38\tfrac{3}{4}x=0 & -14{\cdot}3\ldots, \\
(7) & 16\pm9{\cdot}5+126\tfrac{1}{4}x=0 & -21{\cdot}8\ldots, \\
(8) & 8{\cdot}6\pm10{\cdot}61+219\tfrac{3}{4}x=0 & -56\;\ldots\,.
\end{array}
$$

Nos. (6, 7) require x to be about equal and opposite in the two cases, say, $\delta = 249{\cdot}5$. This would make (8) give $-13{\cdot}5+2{\cdot}61\xi-4(p_1-p_8)=0$. This last case offers some difficulties which we will consider later. For a further approximation we will therefore put $\Delta_2 = 4678+x$, $\Delta'_2 = 4241{\cdot}386+{\cdot}907x$ which give $\delta = 249{\cdot}4933+{\cdot}053x$, and $D_{15} = 20763{\cdot}25+dn$. Then, (p. 366)

$$
\begin{aligned}
d_{15} &= 884207-30{\cdot}50\xi'+30{\cdot}50\,dn\pm{\cdot}5 &&= 189\{4678{\cdot}343-{\cdot}161\xi'+{\cdot}161dn\pm{\cdot}002\} \\
f(2) &= 865448-107{\cdot}26\xi'+16\,p\pm{\cdot}5 &&= 185\{4678{\cdot}098-{\cdot}580\xi'+{\cdot}086p\pm{\cdot}002\}.
\end{aligned}
$$

Hence

$$\cdot 343 - x - \cdot 161\xi' + \cdot 161dn \pm \cdot 002 = 0,$$

$$\cdot 098 - x - \cdot 580\xi' + \cdot 086\,p \pm \cdot 002 = 0.$$

Therefore

$$\left.\begin{aligned} \xi' &= -\cdot 58 - \cdot 38dn + \cdot 20p \pm \cdot 009 \\ x &= \cdot 436 + \cdot 222dn - \cdot 032p \pm \cdot 003 \end{aligned}\right\} \text{A}.$$

Here dn depends on an extrapolated value from a D_3 line, taken because the ν_2 showed a diminished value corresponding to the usual satellite modification of ν_1. If this is correct dn is the observation error on 21858 and is $\cdot 14p$ with $d\lambda = \pm\cdot 03$. This is certainly the most *à priori* probable supposition. If on the contrary hypothesis ν_1 is normal D_{11} is $21552\cdot 65 + \cdot 14p - 786\cdot 45 = 20766\cdot 20$ and $dn = 2\cdot 95 + \cdot 14p$. This gives

$$\xi = -1\cdot 70 - \cdot 05p + \cdot 20p' \pm \cdot 009,$$

$$x = 1\cdot 091 + \cdot 03p - \cdot 03p' \pm \cdot 003.$$

Taking now the case of $D_{18} = 19116$ we have already seen that a change of $y\delta_1$ on the limit, or $2y\delta_1$ is the sequent produce the same change, so that $(y\delta_1)\,D_{18}\,(2y\delta_1)$ is the same as D_{18}, and the D qualification test remains unchanged, although the mantissa is altered by $2y\delta_1 = 174\cdot 8y$.

Supposing this displacement to take place the mantissa of D_{18} is

$$835908 + 124\cdot 8y - 28\cdot 21\xi + 3p \pm \cdot 5 = 197\,\{4243\cdot 188 + \cdot 6330y - \cdot 143\xi + \cdot 015p_8 \pm \cdot 002\}.$$

Here again, as in the case of D_{15}, the question arises whether the value of the normal sequent be taken as that of D_{18} or D_{28} which latter is $\delta_1 = 62\cdot 4$ larger.

The condition then for the exact multiple of Δ'_2 is

$$1\cdot 802 + \cdot 633y - \cdot 907x - \cdot 143\xi + \cdot 015p_8\,(\text{or} + 62\cdot 4/197) = 0.$$

The two cases give

$$\left.\begin{array}{ll} \xi = -\cdot 58, & 1\cdot 505 + \cdot 6330y - \cdot 020p_5 + \cdot 03p' + \cdot 015p_8 \quad (\text{or} + \cdot 316) = 0 \\ \xi = -1\cdot 70. & 1\cdot 063 + \cdot 6330y - \cdot 03\,p_5 + \cdot 03p' + \cdot 015p_8 \quad (\text{or} + \cdot 316) = 0 \end{array}\right\} \text{B}.$$

Neither can be easily satisfied with d_{18}, but with d_{28} the conditions can be satisfied with $y_2 = -3$, -2, for the two cases respectively.*

The remainders in the two cases are $-\cdot 077$ for $\xi = -\cdot 58$ and $+\cdot 113$ for $\xi = -1\cdot 70$.

The list of F lines also give the following which can be used as tests.

(1) (6) $17747\cdot 14 = F_1(2)\,(10\Delta_2)$, *i.e.*, $f(2)$ mantissa $= 195\,\Delta_2$,

(2) (5) $17594\cdot 17 = F_1(2)\,(7\Delta'_2)$,

(3) (2) $17972\cdot 78 = F_1(2)\,(16\Delta'_2 + \Delta_2)$.

---

* See, however, final order below.

The mantissa of (1), using $F(\infty) = 30678{\cdot}93+\xi'$ is

$$912207-112{\cdot}60\xi'+11p_1 \pm {\cdot}5 = 195\,\{4677{\cdot}986-{\cdot}577\xi'+{\cdot}05p_1 \pm {\cdot}002\}.$$

This requires $-{\cdot}014-x-{\cdot}577\xi+{\cdot}05p \pm {\cdot}002 = 0$, or combined with the condition for $f(2)$, $-{\cdot}112+{\cdot}003\xi-{\cdot}086p+{\cdot}05p_1 \pm {\cdot}002 = 0$ and can be satisfied within error limits.

The mantissa of (2) $895149-110{\cdot}63\xi'+11p_2$ which differs from that of f (2) by $29701-3{\cdot}37\xi'+11p_2-16p' = 7\,(4243{\cdot}00-{\cdot}48\xi+1{\cdot}6p-2{\cdot}3p')$.

This requires $1{\cdot}61-{\cdot}907x-{\cdot}48\xi+1{\cdot}6p-2{\cdot}3p' = 0$ easily satisfied for both cases within observation errors.

The mantissa of (3) is $937966-115{\cdot}6\xi'+11p_3$, differing from that of f (2) by $72518-8{\cdot}35\xi+11p_3-16p$. With $\xi = -{\cdot}58$ and $-1{\cdot}70$ this becomes 72523·1... and 72532..., or $17\Delta'_2+1\tfrac{3}{4}\delta-23$... and $17\Delta'_2+1\tfrac{3}{4}\delta-24$... on their respective values of D'_2. The amount 23 is perhaps excessive to be covered by the various possible errors but it just comes within. It may be noted that $17\Delta'_2+1\tfrac{3}{4}\delta = 16\Delta'_2+\Delta_2$. These three data do not decisively distinguish between the two cases. This, however, is not to be unexpected because the two arise from a δ_1 displacement in d_{15}, the sequents in this neighbourhood are such that δ_1 on the limit and $2\delta_1$ on the sequent are nearly equivalent, and the multiples involved 185, 189, 195 are too close to produce contrasts. Incidentally, also, the discussion strengthens the allocation of the lines to the displacements given.

The only further test with our present knowledge is to obtain some independent evidence as to the exact value of the limit, and naturally we turn for this to the mean of the F and **F** series. The series however in Kr is not nearly so well developed as in X. As has been already seen there are only three sets of observed pairs (m = 2, 3, 4) and these give for $F(\infty)$ respectively values of 30674·77, ... 7·27, ... 7·16. Since a displacement of δ_1 produces a change of 2·03 in $F(\infty)$ the first may be due to the fact that the line taken for **F** (2) is really $(3\delta_1)$ **F** (2), when the true mean would be 30677·81. It is natural to seek further as to the existence of summation lines corresponding to our last three examples. The result shows a most remarkable agreement. The sets are shown in the following list together with those obtained from the normal F and **F**.

m.	F.	F (∞).	**F**.	
2	(1) 17321·51	30677·82	(44034·13)	$(-3\delta_1)$ (2) 44028·04
3	(1*n*) 23353·84	30677·27	(2) 38000·71	
4	(2*n*) 26067·66	30677·16	(6) 35286·68	
$F_1(2)\,(7\Delta'_2)$	(5) 17594·17	30677·76	(1) 43761·38	$\mathbf{F}_1(2)\,(7\Delta'_2)$
$F_1(2)\,(10\Delta_2)$	(6) 17747·14	30677·73	(1) 43608·33	$\mathbf{F}_1(2)\,(10\Delta_2)$
$F_1(2)\,(16\Delta'_2+\Delta_2)$	(2) 17972·78	30677·70	(1) 43382·63	$\mathbf{F}_1(2)\,(16\Delta'_2+\Delta_2)$

These are remarkably concordant, especially when it is noted that the F (3, 4) are diffuse lines and not so susceptible of exact measurement as the others. The mean limit is 30677·72 and may be taken as practically correct. That calculated from the series, and used in the preceding discussion is $30678·93+\xi'$. This, therefore, requires a correction of $-1·21$. The equations A would be satisfied by $p=1$, $dn=2$, $\xi=-1·21$. As however this value of ξ is probably correct within ·1 the best value of Δ_2 is obtainable from $f(2)$, viz.,

$$\begin{aligned}\Delta_2 &= 4678·098+·580\,(1·21\pm·1)+·086\times 1\pm·002\\ &= 4678·80\pm·10\\ \Delta'_2 &= 4242·18\pm 06\\ \delta &= 249·536\pm·005.\end{aligned}$$

If the difference between δ as found from ν_1, ν_2 be real and depend on electronic changes as hinted at in the introduction, the changes calculate to 73·94 electrons = 74. In other words, the ν_2 would refer to mass of nucleus +37 electrons and ν_1 to mass of nucleus −37 electrons. Is it merely a curious coincidence that the atomic number of Kr is 38, that of H being 1; all the electrons acting in one way for ν_1 and in the opposite for ν_2? it being remembered that when S lines are emittted one electron at least is absent.

Xenon.— The X diffuse system appears to be a most complicated one. As we shall see later there appears to be a congery of series converging to limits which are collaterals of $S(\infty)=51025$, and connected with these there are again congeries of F series converging to limits collateral to the various d_{1n} sequents or, say, the normal $F(\infty)$. These F series further show the existence of satellites—in other words the f sequence is also subject to slight collateral displacements. This renders their disentanglement a very intricate problem not only in itself, but because it renders the region of the spectrum involved very crowded, with lines close together, with the consequence that coincidences occur which may not refer to real relationship. In fact there are cases where the calculated values of supposed lines of different series are the same within observation errors.* This crowding is also increased by the existence of the allied **F** series referred to above. The complete discussion of all these related series should afford valuable material for arriving at a knowledge of displacement laws. Here it will be sufficient to indicate the nature of the problem and to deal with the material so far as to give confidence in the results as to the assignment of series and especially as to accurate determinations of the oun and the various links.

As vacuum tube spectra approximate to the spark type, the difficulty of drawing definite conclusion from the existence of a triplet separation is again enormously increased by the presence of the link relations which these spectra show. In arc spectra the appearance of a ν_1 or ν_2 separation may always with some certainty be ascribed to the fact that the lines in question are directly connected with series terms.

---

* A case in point is $F_1(5)$ and $F_3(17)$ in the series next considered; also $F_3(13)$ and $F_2(19)$.

Here however no such certain conclusions can be drawn. They may enter as links. Although their true connection may ultimately be definitely settled the doubt as to whether they give true or false scents renders the task of unravelling most bewildering.

In about the region to be expected we find the set

(1) 19880·72	**1778·42**	(5) 21659·14	**815·30**	(10) 22474·44

showing from their separations and order of intensities an indubitable satellite set of D_3 type. Below these come a number of indubitable doublets of D_2 type, and then a number of strong lines of D_{11} type. There are also in this region an extremely large number of lines with separations between 1780—1785 corresponding to the enlarged links already found in the KrD spectrum. This portion of the spectrum is set out together with their separations in the following table :—

(2n) 17903·44	(3) 19632·44	**1780·84**	(3) 21413·28		
1772·81	(7) 19676·25	**1781·23**	(1) 21457·48		
	(1) 19785·84				
	(3) 19815·84				
	(1) 19829·47				
	(1) 19880·72	**1778·42**	(5) 21659·14	**815·30**	(10) 22474·44
	(1) 19942·53	**1775·45**	(10) 21717·98		
	(1) 19959·64	**1772·60**	(1n) 21732·24	**812·77**	(6) 22545·01
(3n) 18238·60	(1) 19989·72	**1780·27**	(6n) 21769·99		
1778·86	(1) 20017·46	**1783·71**	(10) 21801·17		
1783·06	(1) 20021·66	**1779·51**			
(7) 18266·88	(2) 20029·12				
1774·85	(2) 20041·73	**1782·55**	(1n) 21824·28		
	(4) 20080·93	**1784·54**	(1n) 21865·47	**1784·76**	(2) 23650·23
		1786·02	(1n) 21866·95	**1783·28**	
	(1) 20107				

(1) 20305·60		
(6) 20312·70	**1783·72**	(5) 22096·45
(4) 20320·25	**1776·17**	
(1) 20333·22	**1774·92**	(3n) 22108·14
(5) 20443·29		
(5) 20454·88		
(1) 20467·90		
(6) 20470·75		

(7) 20500·13	**1784·68**	(2) 22284·82
1777·11		
(9) 18723·02		
814·13		
(2n) 17908·89		
(3) 20529·92	**1780·35**	(7n) 22310·27
(8) 20559·08	**1785·64**	(<1n) 22344·72
(1) 20581·64		

Before however considering these lines in detail it will be desirable to take here a preliminary discussion which involves a new fact in series relationships, and at the same time will give some reliable data which have a bearing on the present problem of the actual D and F series. The line 20312 and those in its neighbourhood show the following sets of separations, viz. :—

1.	(1) 17615·06	**832·13**	(10) 18447·19	**1865·51**	(6) 20312·70				
2.	(6) 17638·55	**827·92**	(2) 18466·47	**1866·75**	(1) 20333·22				
		1864·74	(5) 19503·29	**829·93**					
3.	(3) 17628·29	**831·70**	(1) 18459·99	**1860·26**	(4) 20320·25		(1) 22180·04		
						1859·79			
4.	(1) 17772·38	**835·41**	(8) 18607·79	**1862·96**	(6) 20470·75				
5.	(3) 19503·29	**829·93**	(1) 20333·22	**830·62**	(2) 21163·84				
				1864·93	(2) 22198·15		(1) 24062·12		(<1) 25925·07
						1863·97		**1862·95**	

In the general survey at the commencement of the investigation a large number of separations of an amount near 1864 was noticed. Suspecting that it indicated the existence of a second type of sharp series, a second smaller separation (for triplets) was looked for, analogous to the $\nu_2 = 815$ of the 1778 set. A further separation was found for a value about 830.* The whole spectrum was searched within the limits 1864 ± 2. The result is shown in the occurrency curve of Plate 2, fig. 3. This curve is unique amongst those hitherto observed in its great height above the pure chance line and also in the steepness of its rise and culmination to a single definite peak. The search brought to light also a very considerable number of long successive chains and of meshes (see *e.g.*, Nos. 2 and 5 above) of the same amount, proving that 1864 enters not only on its original source, as due to a displacement on some fundamental sequence, but also as a link. Now the c link of the normal $\nu_1 = 1778$ is 1872·63 and $c(3\delta) = 1865{\cdot}16$, thus suggesting a possible origin, also a corresponding c link for ν_2, *i.e.*, a separation produced by a second Δ_2 displacement is 833·84, or with a modified $c'(3\delta) = 830{\cdot}70$. These suggest triplets formed by the same Δ_1, Δ_2 on $p(-\Delta_1)$ instead of on p, in other words, series whose limits are $51025 + 1778 = 52803$, 54675·6, 55509·4. But the way in which the separations enter with the suggested D line 20312 indicates that they stand in fundamental relation to it and neither in a linkage relation, nor with the limits named. For in the latter cases it would throw out of gear the whole relation of 20312 with the D set, which some provisional work had seemed to establish. In this work they were considered as part of the D system through 20312 regarded as a D_1 (1) line with 18447, 17615 as satellites. In this case the d_1 sequence is of the order $51025 - 20312 = 30713$. As against this idea is, of course, the greater intensity of 18447, the supposed satellite over that of the D_{11}, and also the fact that no normal ν_1

---

* As a fact, however, there are several others depending on oun multiples also present, but which at present we need not deal with.

separation occurs with 18447 or ν_1, ν_2 with 17615. But provisionally that was set aside for the moment. If they represented a special D set, the separations ought to reappear in a triplet series of the F type, and in the reverse order. From the sets already excerpted the lines (8) 18607·79, (10) 23915·72, (5) 26365·19 appeared to have all the signs (RYDBERG'S tables) of belonging to one series. The formula calculated from them brought to light a whole long series of observed lines. The limit found was 30724, close to the value already found (30713) as of the order of magnitude to be expected. This so far supported the supposition of the D relation, but there also came to light another result of evident importance in general theory—viz., the **F** series already referred to. The ordinary form of a series is one in which successive lines obey a formula of the type $A - \phi(m)$. In this case we find series associated with it whose successive lines are given by $A + \phi(m)$. This holds for each of the triplet sets, so that the complete series are given by $A \pm \phi m$, $B \pm \phi(m)$, $C \pm \phi(m)$, where $B = A + 1864$, $C = A + 1864 + 830$. Quite apart from the importance of this fact in the theory of spectral series the phenomenon is of special use in calculating the various constants on which the series depends. For instance the sum of the wave-numbers of two corresponding lines gives 2A, 2B, 2C, thus determining the values of the limits quite independently of the nature of the series formula used. Moreover, the displacements which so frequently occur in the F and D series in the sequence term introduces uncertainties. This happens in two ways. First through the modified ν_1 values in which it is not always possible to say whether the displacement is produced in the D_1 or the D_2 line. Secondly because the typical line in any order is often wanting and only appears with a very large displacement of multiples of Δ_2 on the sequence term. This effect, however, provided it occurs for both sets (F, **F**), does not influence the values of A, B, C thus determined. Cases in point are the Kr sets $F(2)(7\Delta'_2)$, $\mathbf{F}(2)(10\Delta_2)$, $\mathbf{F}_1(2)(16\Delta'_2+\Delta_2)$ given on p. 380. In consequence it is possible to determine the separations $B-A$, $C-B$ independently of satellite or other displacements. That such sequence displacements occur in these 1864 series is shown by separations which deviate from the normal by more than observation errors.

But, further, the difference of two corresponding F and **F** lines, say $\mathbf{F}_1 - F_1$, $\mathbf{F}_2 - F_2$, $F_3 - \mathbf{F}_3$, should each give $2f(m)$, if as is the normal rule the sequence term is the same for each line of a triplet. When however—as we have seen in Kr, and shall find even more markedly in X—there are displacements in $f(m)$ for successive lines in a triplet, these differences will not be the same, and the observed separations will vary from the normal values. For instance, suppose $f(m)$ becomes $f(m)-x$ for the second set, and $f(m)-y$ for the third. The lines are $A \pm f(m)$, $B \pm (f(m)-x)$... . The values of A, B, C calculated from the sums are not affected, and the real values of the separations given by $A-B$, $C-B$ are not affected although the observed separations are $\nu + x$, $\nu' + y$ and $\nu - x$, $\nu' - y$. In some cases we shall find evidence from close lines with different x or y—but the results are quite definite. If, however, in the corresponding terms of the F and **F** lines the f are different, then the value calculated from their sum shows a change from the normal limit. The effect shows itself at once and the interpretation is less certain. It is possible that where such an affect appears it may not be real, but due to the existence of the two close lines just referred to, of which one in each set is too faint to have been observed. Thus if the displacements are x_1, x_2 instead of finding $B-f+x_1$, $B-f+x_2$, $B+f-x_1$, $B+f-x_2$, the 2nd and 3rd, or the 1st and 4th may not have been observed and we should be led to a wrong conclusion by taking, say, the 1st and 4th as corresponding lines. There are cases of this kind and also where one only is absent—*i.e.*, we find one close doublet for one of the F or **F** lines.

The lines composing the series are given in the table below. The limit calculated from the first three F_1 lines was found to be 30724·28 ± 1·80, the uncertainty being due to supposed maximum observation errors of ±·05A in each line. The later discussion of the $\frac{1}{2}(F+\mathbf{F})$ rule will show that the limit should be very close to 30725·26 with an error probably < ·3. The formula was recalculated with this limit by supposing the three standard lines to be in error by −·02, + ·02, −·02, *i.e.*, by half their supposed maximum possible errors. The formulæ for the F_1 and $\mathbf{F}_1$ series then become

$$n = 30725{\cdot}26 \mp N \Big/ \left\{ m + 1{\cdot}022746 - \frac{{\cdot}028705}{m} \right\}^2.$$

List of F and **F** Lines.

In each order the first line of numbers gives the F set, the second the **F** set. Between these are entered the mean values of the F and **F** which give the corresponding limits. When the values are deduced by methods explained in the notes they are enclosed in ( ), when calculated and not observed in [ ].

= 1	(3010·35)	**1864·64**	(4875·09)	**830**	(5705·00)
	30725·77		**32590·35**		**33420·35**
	(58441·20)	**1864·50**	(60305·70)	**830**	(61135·70)
= 2	(8) 18607·79	**1862·96**	(6) 20470·75	**829·46**	(<1) 21300·21
	30726·05		**32589·09**		
	(4) 31102·16.*e.v**	**1863·17**	(<1) 32965·33.*e.v*		
= 3	(10) 23915·72	**1863·92**	(8) 25779·64	**829·69**	(1) 26609·33
	30725·15		**32589·28**		**33418·35**
	(2) 37534·58	**1864·35**	(δ_1) (2) 39404·36	**828·44**	(1) 40227·37
= 4	(5) 26365·19	**1864·61**	*u*.(1) 32362·98†	**829·67**	(29059·47)
	30724·58		**32589·96**		**33420·07**
	(35083·97)	**1866·16**	(36950·13)	**831·54**	(37781·67)
= 5	(5) 27696·15	**1865·51**	(1) 29561·66	**828·72**	(1) 26257·20.*u*
	30725·48		**32589·38**		**33419·32**
	(1) 33754·81	**1862·29**	(<1) 35617·10	**831·17**	($-\delta_1$) (1*n*) 36442·62

* *More probably* $\mathbf{F}_1$ (16).

† *Or* $\mathbf{F}_2$ (21).

LIST of F and **F** Lines (continued).

$m = 6$	$(2\delta_1)$ (1) 28508·55	**1864·13**	$(-\delta_1)$ (<1) 30357·31	**829·19**	$(-2\delta_1)$ (<1) 31180·63
	30725·29		**32589·86**		**33417·27**
	(1) 32951·97	$\mathbf{186^{3\cdot 81}_{5\cdot 76}}$	(3) 34815·79 (5) 34817·73	**826·83**	(1) 35642·62
$m = 7$	(1) 29019·32	**1864·99**	(5) 30884·31	**832·82**	(5) 31717·13
	30724·37		**32588·57**		**33418·13**
	(2) 32429·42	**1863·42**	(1) 34292·84	**826·30**	$(2n)$ 35119·14
$m = 8$	[29377·00]	[**1865·39**]	(1) 31242·39	**829·13**	$(-\delta)$ (1) 32048·92
	30725·60		**32590·19**		**33418·91**
	(1) 32074·21	**1863·78**	(2) 33937·99	**828·31**	(δ) (2) 34788·90
$m = 9$	(1) 29629·10	**1865·33**	(31494·43)	**831·94**	(2) 32326·37
	30724·64		**32589·98**		**33418·72**
	(2) 31820·18*	**1865·36**	$(-2\delta_1)$ (3) 33674·68	**825·54**	(4) 34511·08
$m = 10$	$(-\delta)$ (1) 29802·26	**1863·53**	$(-\delta_1)$ (3) 31680·16	**828·83**	$(2\delta_1)$ (4) 32525·72
	30725·20		**32589·74**		**33418·29**
	(2) 31628·35	**1865·54**	(4) 33493·89	**830·27**	$(5n)$ 34324·16
$m = 11$	(29967·00)	**1863·04**	$(-2\delta_1)$ (2) 31820·04*	**832·46**	$(-2\delta_1)$ (3) 32652·20
	30725·18		**32590·25**		**33420·21**
	(5) 31483·37	**1866·08**	(2) 33349·45	**827·08**	(<1) 34176·93
$m = 12$	$(2\delta_1)$ (6) 30091·12	**1864·59**	$(-2\delta_1)$ $(2n)$ 31934·91	**830·81**	$(-2\delta_1)$ (1) 32787·85
	30725·25		**32589·48**		**33419·31**
	(1) 31369·31	**1863·88**	$(-2\delta_1)$ (3) 33222·37†	**828 85**	(1) 34062·04
$m = 13$	$(-2\delta_1)$ (1) 30157·45	**1862·96**	(1) 32030·34	**831·00**	(3) 32861·34‡
	30725·48		**32589·18**		**33419·61**
	$(-2\delta_1)$ (3) 31273·55	**1864·64**	(1) 33148·03§	**829·86**	(4) 33977·89
$m = 14$	(4) 30239·16	**1862·57**	(δ_1) (<1) 32107·16		
	30724·58		**32588·06**		
	$(2\delta_1)$ 31219·95	**1864·39**	$(-3\delta_1)$ (5) 33059·49		
$m = 15$	[30297·01]	[**1865·42**]	$(-\delta_1)$ $(2n)$ 32166·86	**828·06**	$(-\delta_1)$ (2) 32994·92
	[**30725·13**]				**33418·33**
	$(-\delta_1)$ (4) 31148·28		**1864 + 829·46**		$(-\delta_1)$ (1) 33841·74
$m = 16$	(1) 30348·83	**1864·97**	(<1) 32213·80		
	30725·49		**32589·56**		
	(4) 31102·16	**1863·17**	(<1) 32265·33		(δ_1) (5) 33799·54
$m = 17$	$(-\delta_1)$ (<1) 30382·49	**1861·03**	(3) 32248·49 ?		
	30724·89				
	$(-3\delta_1)$ (<1) 31047·46		**1864 + 828·48**		(1) 33754·81
$m = 18$	(1) 30421·95	**1863·31**	$(-\delta_1)$ (1) 32279·83	**829·68**	$(-2\delta_1)$ (<1) 33103·59 and
					$(2\delta_1)$ (<1) 33126·29
	30725·06		**32588·13**		**33418·27**
	(δ_1) (1) 31033·15	**1862·42**	(δ) (1) 32912·71		$(-\delta_1)$ (3) 33715·67 and
					(δ_1) (3) 33724·54

* *See* $\mathbf{F}_2$ (11), *is more probably* $(-2\delta_1)$ (1) 31810·86.

† *This line is numerically* $(2\delta_1)$ $\mathbf{F}_2$ (12) *and* $(-2\delta_1)$ F_3 (22).

‡ F_3 (13) *and* $\mathbf{F}_2$ (19) *have same value.*

§ $\mathbf{F}_2$ (13) *and* F_3 (19) *have same value.*

LIST of F and **F** Lines (continued).

= 19	$(-2\delta_1)$ (<1) 30444·37	**1861·79**	(1*n*) 32316·13	**831·90**	(1) 33148·03†
	30725·42		**32588·93**		**33420·61**
	(2) 30996·51	**1864·83**	(3) 32861·34*	**831·84**	(3) 33693·18
= 20	$(-2\delta_1)$ (1) 30466·53	**1864·01**	(5) 32340·48	**830·69**	(δ_1) (2) 33165·52
	30525·60		**32589·36**		**33418·76**
	$(-5\delta_1)$ (1) 30949·88‡	**1863·52**	(2) 32838·25	**828·73**	(1) 33666·98
= 21	(δ_1) (3) 30503·24	**1864·71**	(1) 32362·98	**828·68**	$(-\delta_1)$ (<1) 33186·10 an
					(δ_1) (<1) 33197·22
	30724·07		**32589·61**		**33418·10**
	(1) 30949·88	**1866·37**	(3) 32816·25	**828·29**	(2) 33644·54
= 22	$(-\delta_1)$ (3) 30511·25	**1865·32**	$(3\delta_1)$ (<1) 32397·90§	**829·46**	$(2\delta_1)$ (3) 33222·37‖
	30725·36		**32588·60**		**33421·74**
	(5) 30933·51	**1862·19**	(1) 32795·70	**830·97**	(δ_1) (1) 33632·32
= 23	[30534·18]	**1863·72**	(<1) 32397·90§		
	30725·57 + $\frac{1}{2}\xi$		**32590·16**		
	(<1) 30916·97	**1685·45**	(δ_1) (1) 32787·85		
= 24	$(+2\delta_1)$ (3) 30559·83				
	30725·25				
	$(-\delta_1)$ (4) 30895·10 and	**1864·66**	(1) 32765·29§		
	(δ_1) (1) 30906·17				
= 25	$(-3\delta_1)$ (<1) 30576·46¶	**1864·57**	(1) 32426·15	**829·45**	(δ) (3) 33278·20††
	30725·68		**32590·24**		
	$(-\delta_1)$ (5) 30884·81**	**1864·55**	$(2\delta_1)$ (1) 32765·29§		
= 26	(<1) 30576·46	**1864·16**	$(-2\delta_1)$ (2) 32429·42§§	**829·83**	(1) 33270·45
			and $(2\delta_1)$ (3) 32451·83		
	30725·37		**32589·73**		**33420·50**
	(δ_1) (1) 30879·26	**1864·55**	$(-2\delta_1)$ (4) 32727·98	**831·70**	$(2\delta_1)$ (<1) 33581·84
= 27	$(-\delta_1)$ (5) 30580·76		**1863·50 + 828·94**		(3) 33278·20
	30725·87				
	$(-\delta_1)$ (6) 30861·05		**1864·5 + 829·46**		$(-\delta_1)$ (6) 33553·33
= 28	(1) 30595·63	**1864·11**	(4) 32459·74	**830·43**	$(2\delta_1)$ (3) 33301·47
			32589·87		
	[ ]		(2) 32719·41	**828·43**	$(-3\delta_1)$ (2) 33530·95‡‡
= 29	(1) 30607·90		**1865·65 + 837·92**		(3) 33301·47
	30725·67		**32590·00**		**33419·00**
	(1) 30843·45	**1863·01**	(3) 32706·46	**830·14**	(2) 33530·95‡‡
= 30	$(2\delta_1)$ (4) 30621·49		**1864 + 829·36**		(3) 33304·91
	[30839·22]	[**1861·57**]	(<1) 32700·79	**830·16**	(2) 33530·95‡‡

* $\mathbf{F}_3$ (13) *and* $\mathbf{F}_2$ (19) *have same value.*

† $\mathbf{F}_2$ (13) *and* F_3 (19) *have same value.*

‡ *The line is* $\mathbf{F}_1$ (21); $\mathbf{F}_1$ (20) *may be* (4) 30976·73.

§ *Coincidences.*

‖ *This line is numerically* $(2\delta_1)$ $\mathbf{F}_2$ (12) *and* $(-2\delta_1)$ F_3 (22).

¶ *This line is also* F_1 (26).

** *This line is* F_2 (7).

†† *This line is* F_3 (27).

‡‡ *Too close to settle.*

§§ *This line is also* $\mathbf{F}_2$ (7).

The lines of the series seem to be exceptionally numerous. The results of the examination up to $m = 30$ are given in the table and the notes thereto. There are certain lacunæ—especially for $m = 4$. In these cases however corresponding displaced sets are in general observed, and naturally with large values of m this effect is more frequent. In certain cases where a set is absent a parallel set is observed linked to the normal type. This is the case for instance in $m = 4$.

The question naturally arises whether lines exist for $m = 1$. If so the formula gives a triplet with the first line at $n = 3142$, far in the ultra-red. In other spectra these values extrapolated for $m = 1$ differ considerably, often by several hundreds, from the correct ones. We can only conclude that if there are sets based on $m = 1$ they must be such that F_1 (1) must be in the neighbourhood of 3100. The matter can only be settled therefore by other considerations which must depend—with our present knowledge at least—either on sounding or on the presence of combination lines in the observed region. The evidence for such a triplet is given below in the notes to the list of lines. The value of F_1 (1) found is 3010·35 corresponding to a wave-length *in vacuo* of 33218·7 A.U. The mantissa of $3010+dn$ with the limit $30725\cdot26+\xi$ is $989285+35\cdot9\,(dn-\xi) = 90\,\{10998\cdot8-\cdot4\xi+\cdot4dn\}-611 = 90\Delta_2-\delta$. The uncertainty in Δ_2 as found from ν_2 is too large to settle the exact value of this with so large a multiple as 90, but the fact as it stands that the mantissa differs from a multiple of Δ_2 by only a few ouns is what is to be expected if the series belongs to the F type, and so far certainly supports the more direct evidence given below for the existence of the set depending on $m = 1$. With the value of Δ_2 found below $dn = -1\cdot5$.

A glance at the list will show that the separations observed in the second and third orders of F are less than the normal values. This points to a satellite effect. The values of ν_1 are 1862·96, 1863·92 which show deficits of 1·54, ·58 from the true value as indicated by the occurrency curve. Now a displacement by one oun produces a change of 1·25 in $m = 2$, and ·50 in $m = 3$. The deviation is then completely explained by the supposition of the existence of the satellite effect depending on δ_1. The ν_2 show similar deficits, which may possibly be due to observation errors. We should expect to find a similar effect (not necessarily the same multiple) in the order $m = 1$. In this order the oun produces a change of 4·25.

For sounders and for link evidence the data have been restricted to *e*, *u*, *v* links only. If we may judge from the examples of Ag and Au, the F and D linkages show a preponderance of the *a*, *b*, *c*, *d* links, and no doubt fuller evidence might have been adduced by using them, but it was necessary to set limits to the work, as well as to this communication which is long as it stands. But as examples we may give some *d* links belonging to the orders 2, 3 of F_1. The value of *d* is 1973·94. For $m = 2$ the lines (1) 20581·64, **1864·85** (3) 22446·49, are 1973·85, above F_1 and $F_1+\nu_1$, or the real F_2. The two lines (5) 23268·21, (2) 82·67 are respectively $\mp 6\delta_1$ displacements of 23275·44 which is 1875·23 (or *c*) above the observed F_3 or 1873·15 above the normal value $F_1+\nu_1+\nu_2$. For $m = 3$ we find

			(2) 18466·47				
			7313·17				
(1) 28341·62 **4425·90** (4) 28047·34 **4131·62**	F_1	**1863·92**	F_2	**829·69**	F_3	**4426·43**	(2) 31035·76
	1975·86		**1974·70**		**1970·10**		
	(4) 25891·58	**1862·76**	(2) 27754·34	**825·09**	(4) 28579·43	**1864·94** →	

Here F_1, 25891, 27754 form a series inequality with $d+2$, ν_1-2; F_1 forms a parallel inequality with 28341 and 28047 with $u-2$, $v-2$; 18466, F_2, 27754 a series inequality with d, e; and F_3 a parallel inequality with $v-2$, $d-2$.

As a rule in the list only displaced lines are dealt with and only occasionally linked ones. The evidence is strong for the existence of possible series with limits 30725 ($\pm x\delta_1$)—especially $x = 2$—and that for each order most of the energy goes to a line with one of these limits, not necessarily the same for different orders nor for the same triplet set, nor for the same F or **F** type.

Notes.—A displacement δ_1 produces 4·97 on $F_1(\infty)$, 5·43 on $F_2(\infty)$, and 5·65 on $F_3(\infty)$.

$m = 1$. The shortest sounder to reach from the calculated F (1) to the observed region is $2e$, giving values in the neighbourhood of 17700. In this neighbourhood we find a triplet (No. 2 of the sets on p. 383)

(6) 17638·55 **1864·64** (3) 19503·29 **829·93** (1) 20333·22,

which in separation and order of intensity corresponds precisely to the values required. With the normal value of e they indicate lines with wave-numbers at 3010·35, 4875·09, 5705·02. In connection with the question of satellite effect it may be worth noting that lines in the neighbourhood of the above may be arranged with them thus

(6) 17638	**1864·64**	(3) 19503	**829·93**	(1) 20333
		12·52		
....		(3) 19515·81		
25·53				
(5) 17664·07				

To complete the set there should be a line at 17651·31, but it was not observed by BALY. A sequence displacement in $f(1)$ of $3\delta_1$ produces a change of 12·76 and of $6\delta_1$ of 25·53. These are practically exact and correspond to the diffuse satellite arrangement with this difference, that the triplet appears as a main set in which the first line is strongest. They depend on $3\delta_1$, $6\delta_1$ displacements, so that the mantissa of the doublet set, that of 19515, is exceedingly close to $90\Delta_2$, viz., $90 \times 10997·17$.

With $F_1(1) = 3010$ should go $\mathbf{F}_1(1) = 58440·17$. This is outside the observed region on the other side, in the ultra-violet, and requires even larger sounders than F (1). The treble link $3e = 21942·30$ requires a line at 36497·87. There is no line here but (1) 36493·82 is less by 5, *i.e.* = $(\delta_1)\,\mathbf{F}(1)$. We can arrange lines to go with it as a satellite set, viz.,

[36498·90]	**1864·50**	[38363·40]	**830**	(1) 39193·40
		17·64		**17·66**
....		(2) 38345·76	**829·98**	(2) 39175·74
25·45		**26·31**		
(3) 36473·45	**1863·64**	(1) 38337·09		

Here we find the same $6\delta_1$ for $\mathbf{F}_1$ and $\mathbf{F}_2$, whilst in F we had it only for F_1. The set with separation 17 show no $\mathbf{F}_1$, and again is numerically an exact oun multiple displacement, viz., δ which gives 17·0. The line found close to the expected—$(\delta_1)\,\mathbf{F}_1$ above—forms part of a chain

(1) 32765·29 **1864·47** (3) 34629·76 **1864·06** (1) 36493·82

It may be the representative corresponding to the satellite displacement δ_1 observed in F for $m = 2, 3$.

Using links $2e+u$ = 18761·38 and $2e+v$ = 19056·2 we should find lines at 39678·79 and 39383·97 with the corresponding lines to $\mathbf{F}_2$, $\mathbf{F}_3$ beyond the observed region. As a fact, we find lines at (3*n*) 39683·95 and (1) 39386·98. The first corresponding to a $-\delta_1$ displacement should give $2e.v.\mathbf{F}_1(1)$ = 39679·70. The second would differ from the same displacement by 3 and is therefore inadmissible. It should be noticed that both the $3e$ and $2e+u$ sounders give lines larger by unity than the expected value. This may, of course, be due to combined errors in the sounders and observation errors, but the **F** are entered in the list as if the sounders are correct. As none of the F (1) or **F** (1) lines can be observed, their means can have no weight for an accurate determination of the limit.

$m = 2$. The **F** lines are in the ultra-violet, and should be, with the limit chosen, at 42842·73, 44707·23. With sounder $e+v$ = 11742·10 we find the set (4) 31102·16 **1863·17** (<1) 32965·33 with a small separation corresponding to that for F_1. It gives $\mathbf{F}_1$ = 42844·26 and $\mathbf{F}_2$ = 44707·43. But the set appears to be really $\mathbf{F}_1$ (16), $\mathbf{F}_2$ (16).

$m = 3$. A displacement δ_1 in $\mathbf{F}_2(\infty)$ produces 5·42. The $\mathbf{F}_2$ line is not observed but (2) 39404 is practically an exact $(-\delta_1)\,\mathbf{F}_2$. In further illustration of this we find (2) 37431·03 linked to it by $-d$ and 5·61 (*i.e.* another δ_1) ahead in the remainder of the triplet (1) 37436·64 **828·13** (1*n*) 38264·77. The lines in question are $(-\delta_1)\,\mathbf{F}_2$; $d.(-2\delta_1)\,\mathbf{F}_2$; $d.(-2\delta_1)\,\mathbf{F}_3$.

$m = 4$. $\mathbf{F}_2$ is (1) 32362·98 − u = 28229·80 which, as referred to above, is also (1) 26257·20 + 1972·60 close to d link. For F_3, $(-\delta_1)$ (4) 29053·89 = 29059·54 or $(2\delta_1)$ (1*n*) 29071·03 = 29059·73 or $(3\delta_1)$ (1) 29076·10 = 29059·15. Take the mean 29059·47 as $\mathbf{F}_3$. Several other lines in this neighbourhood show indications of displacement by multiples of δ_1, *i.e.* that the energy proper to $\mathbf{F}_3$ has gone into a number of collaterals. None of the direct **F** are observed but there appears a parallel set displaced δ, viz., (1) 35064·09 **1864·34** (1) 36928·45 with (1) 37742·12 due to a further $3\delta_1$ displacement. The **F** lines are entered as due to these.

Both separations are about 1·6 too large and further $\mathbf{F}_1$ is about the same amount too small. Now $2\delta_1$ on the sequence term produces a change of 1·40. This would indicate that with the δ displacement on the limit concomitant displacements of $2\delta_1$ occur in the sequence so that the observed lines are $(\delta)\,\mathbf{F}_1\,(-2\delta_1)$, $(\delta)\,\mathbf{F}_2$, $(7\delta_1)\,\mathbf{F}_3\,(2\delta_1)$. The whole set again is remarkable for connection with parallel sets separated by the normal $\nu_1 = 1778$.

$m = 5$. There are observed lines for $\mathrm{F}_{1\cdot 2}$, F_3 is entered as depending on a u link. F_2 is too large by about one unit, but $\mathrm{F}_3 - \mathrm{F}_1$ is normal. This order affords good evidence for the existence of the displaced sets. Consider the following system of lines:—

		1863·57	(3) 29540·09	**830·59**	
(δ) F	(1) 27676·52		**1·84**		(1*n*) 30370·68
		1865·41	(2) 29541·93	**828·75**	
	19·63				**11·81**
($2\delta_1$) F					(<1) 30382·49
F	(5) 27696·15	**1865·51**	(1) 29561·66		
	15·51				**22·26**
($-3\delta_1$) F	(1) 27711·66	**1863·73**	(3) 29575·39	**829·36**	(5) 30404·75
($-\delta$) F	(5) 27714·80	**1869·68**	(2) 29584·48		

Here 1·84 is an exact 3δ displacement in the sequent. The mesh shows series inequalities with the $\nu_1 + \nu_2 =$ normal values. The two lines in question are clearly $\pm 6\delta_1$ displacements on a normal line $(\delta)\,\mathrm{F}_2$. On $\mathrm{F}_1(\infty)$, δ gives 19·88, and $3\delta_1$ 14·91; on F_3, $2\delta_1$ gives 11·30 and δ 22·60. These show how closely all the conditions of the allocations are satisfied. Further, it shows how 29561 has its excess value and that the normal sequent should be the same as in $(\delta)\,\mathrm{F}_2$, *i.e.*, 9 less, thus making its separation with $\mathrm{F} = 1864·61$. The same sequent change is shown by $\mathbf{F}_2$. Both give it as $(6\delta_1)f$.

$m = 6$. Again with an even order the F lines do not appear, but there are apparent also a congery of displaced lines analogous to that in F (5). The lines given in the list give wave-numbers 28498·61, 30362·74, 31191·93. On the contrary, **F** lines are observed although $\mathbf{F}_3$ has probably been displaced. These also show evidence of displaced sets, *e.g.*, (1*n*) 32972·50 **1864·28** (<1) 34836·78 is 20·53 ahead of $\mathbf{F}_1$, and δ on the limit gives 19·88.

$m = 7$. The values of $\mathrm{F}_1(\infty)$, $\mathrm{F}_2(\infty)$ as deduced from the means are clearly too small. $\mathrm{F}_1(7)$ is very close to the calculated value, so that if any error has been made it is probably due to the **F** which should be about 1·8 larger, and suggests a close doublet, *i.e.* a small sequence displacement as in the preceding sets. As supporting this there is a line (1) 32426·15 which as $(\delta_1)\,\mathrm{F}_1$ would give 32431·12 making with F_1 the limit 30725·22. This corresponds clearly to the normal value. A similar displacement is also found in F_1 in the line (4) 29024·43, which as $(-\delta_1)\,\mathrm{F}_1$ gives 29019·46 for F_1. It should be noted that the energy of F_1 has passed chiefly to the displaced line, whilst in F_2 most of its energy remains with it and a fraction passes to the displaced line. This probably means that only a small number of the normal F_2 configurations are broken up, whilst most of the F_1 are. F_3 as (5) 31717·13 gives ν_2 too large. This line and (4) 31705·47 are separated by 11·66 or a $2\delta_1$ displacement, so that there is a concomitant sequence displacement. A similar effect is shown in $\mathbf{F}_3$ with two lines (2*n*) 35119·14, (2) 35126·05. The lines entered appear correct for they give the normal limit, but their half difference shows a displacement in the $f(7)$ sequent. The normal line would appear to be given by $(-2\delta_1)\,\mathbf{F}_3 = 35136·05$ making $\mathbf{F}_3 = 35124·78$ with $\nu_1 = 831·94$.

$m = 8$. No line is found for the calculated 29377·00, or 77·23 if we allow the same O − C as for $m = 7$. The lines (3) 29368·41 as $(2\delta_1)\,\mathrm{F}_1$ and (1) 29403·29 as $(-5\delta_1)\,\mathrm{F}$ give respectively 29378·34 and 78·45, which are larger than should be expected. The calculated value has been taken as correct. Also the lines (1) 32048·92 as $(5\delta_1)\,\mathbf{F}_1$ and (4) 32098·20 as $(-5\delta_1)\,\mathbf{F}_1$ give respectively 32073·76 and 73·36 or a mean of 73·56. But 32048 is also $(\delta) F_3$. This is not a mere list coincidence. As a fact $(\delta) \mathbf{F}_3$ and $(5\delta_1) \mathbf{F}_1$ are very nearly equal, and if both existed would show as a double line too close to have been resolved. The second line has a separation 1864·23 to (2) 33962·43, and its deduced $\mathbf{F}_1$ makes with the calculated F_1 the limit 30725·18 very close to the definitive value found below. For $\mathbf{F}_2$ 33937·99 is supported by $(-2\delta_1) \mathbf{F}_2 = (1)\ 33948{\cdot}71$ giving $\mathbf{F}_2 = 33937{\cdot}85$, but the $F_2(\infty)$ is large and 31242 shows a separation with F_1 of 1865·39 also large. Also (1) 31247·07 as $(-\delta_1) F_2$ gives $F_2 = 31241{\cdot}64$ which makes $F_2(\infty) = 32589{\cdot}82$ and the separation from $F_1 = 1864{\cdot}64$ both improved. On the other hand $(-2\delta_1) F_2 = (4)\ 31253{\cdot}32$ gives $F_2 = 31242{\cdot}46$ precisely the line observed. These small differences depend partly on observation errors and sequence or satellite displacements. In the case of F_3 and $\mathbf{F}_3$ the equally and oppositely displaced lines give the same mean as the lines calculated from them.

$m = 9$. There seem considerable displacements in the sequences here. The calculated values for the first lines are 29632·80 and 31817·72. They are not observed, but the corresponding F_3, $\mathbf{F}_3$ lines are. There are two near observed lines (1) 29629·10 and (2) 31820·18 which give the mean 30724·64, which is small, but (1) 31810·86 = $(2\delta_1) \mathbf{F}_1$ would give 31820·79 and the limit 30724·94, close to the normal value. 31820·18 is then $(2\delta_1) \mathbf{F}_2(11)$. With (5) 31483·37 as $(2\delta_1) \mathbf{F}_2$ and (1*n*) 31505·49 as $(-2\delta_1) \mathbf{F}_2$ we get respectively $F_2 = 31494{\cdot}23$ and 94·63. The mean is entered, and a similar $-2\delta_1$ displacement gives $\mathbf{F}_2$ as entered. The normal third lines are observed. Probably the $\mathbf{F}_3$ having the same sequent as the $\mathbf{F}_2$ adopted should be that given by $(-2\delta_1) \mathbf{F}_3 = (4)\ 34527{\cdot}64$ or $\mathbf{F}_3 = 34516{\cdot}37$ with $\nu_2 = 830{\cdot}83$.

$m = 10$. The allocation seems satisfactory. The limits also are very close to the correct, but the different triplet separations show that the successive sequents suffer displacement, but the same in each F, **F**.

$m = 11$. The calculated F_1 is 29966·21. With (1) 29982·13 = $(-\delta_1) F_1$, $F_1 = 29967{\cdot}22$. Moreover the last has links $e = 7314{\cdot}34$ to (3) 22652·88 and $u = 4133{\cdot}20$ to (1*n*) 25854·02 in very striking agreement. The value as calculated with normal *e* is entered. With the lines as entered it is seen that the means of the corresponding separations for the two series are both normal, although the individuals are abnormal. This shows that both corresponding lines have the same limit, and the same sequent, but that the latter shows a displacement from the normal value for the F_2 set. This is supported also by the fact that there are a number of close lines to $\mathbf{F}_2$. For instance, (2*n*) 33332·22 and (2) 33330·00 as $(3\delta_1) \mathbf{F}_2$ give respectively 33348·51 and 6·29 for $\mathbf{F}_2$. They are probably all $\mathbf{F}_2$ lines showing sequence displacements. The first gives the triplet separations 1865·14, 828·42 and limit 32589·77, the second 1862·92, 830·64 and limit 32588·66. In other words, the first gives $\mathbf{F}_2$, with same sequent as in $\mathbf{F}_1$ and F_1, the second as in F_2.

$m = 12$. Note the good agreement—the same $(-2\delta_1)$ displaced limit for F_2. $\mathbf{F}_2$ and same $2\delta_1$ for F_1 and $\mathbf{F}_3$.

$m = 13$. The two displaced sets give $F_1 = 30167{\cdot}38$, $\mathbf{F}_1 = 31283{\cdot}59$. The calculated $F_1 = 30166{\cdot}40$.

$m = 14$ *to* 30. It is remarkable how the series seems to persist to high order. It may be said that this is only apparently so, because in this region the spectrum is so crowded with lines that it is necessarily possible to select sets near the calculated values. But in truth the reason of the crowding is because of the series. The F and **F** lines crowd up together on either side of the three limits, and at the same time there are different sets of limits depending on the $2\delta_1$ displacements. The spectrum has not been examined beyond $m = 30$, and from 14 to 30 the list indicates an allocation without further specification. There is, however, much evidence not adduced here to indicate actual cases where sequence displacement occurs. The calculated values for F_1 from $m = 14$ to 30 are 30238·21, 30297·01, 30345·73, 30386·58, 30421·17, 30450·68, 30476·20, 30498·14, 30517·35, 30534·18, 30549·11, 30562·32, 30574·08, 30584·61, 30594·07, 30602·60, 30610·32. The deviation from the calculated values for $F_1(29)$ and $\mathbf{F}_1(29)$, which, however, gives the correct limit, shows that the sequent $f(29)$ receives a large displacement value, so large indeed as to totally alter its mantissa. The set must be doubtful. The whole set are more likely to be $(\delta_1)\mathrm{F}(30)$ and $(-\delta_1)\mathrm{F}(30)$. In fact, in this neighbourhood, the difference of two successive orders in a series is comparable with the change produced by a δ_1 displacement in the limit, and so introduces some uncertainty in allocation. It will be noticed also that, in a few cases, the same line is adduced to fit two cases, which can only happen if a line happens to be a close doublet, an unlikely supposition to happen often.

We are now in a position to determine the limits with considerable accuracy. Taking the average means where they are deduced from actual or displaced actual values, we find the limits come to 30725·340, 32589·443, 33419·079. In the first attack on the problem values of unobserved lines were deduced from observed linked lines. The corresponding mean values for the limits then found had for the last digits 5·292, 9·161, 8·918, very close to those determined from the displacements. The individual deviations from the mean are quite small for $\mathrm{F}_1(\infty)$, considerably smaller than for the others. It is, therefore, the more reliable. The mean deviation in magnitude is ·28 and the maximum is −·97 in $m = 7$. We may take it therefore that the true value of the F_1 limit is 30725·30 within a few decimals. The separations given by the deduced limits above are 1864·10 and 829·64.

These very accurate values afford a means of testing as to their source. If the limit were known to be a single number there could be no doubt as to its belonging to the *d* sequences, or as to the series being of the F type. But there is just the possibility that it may be a composite number, comprising one or more links—say, *p* or *s* terms—and that the separations may be due to oun displacements in one of them. The suspicion that this may be the case is aroused by the fact that the triplet 17615, 18447, 20312, which would be the origin of the *d* or $\mathrm{F}(\infty)$ term, and in which therefore the first two lines should behave as satellites do not show complete sets with the separation 1778, 815, as they should do if normal satellites. Moreover, the intensity order with the middle line much more intense than the other is not normal. There is no test for the composite nature of 30725, but if it be really so, the most probable source would be $p = \mathrm{S}(\infty)$, or some near collateral of this. We will therefore test this as $51025{\cdot}26 + \xi$, where ξ may be considerable, so as to include near collaterals, and also test 30725 as a *d* sequent. We will take the latter first.

At the start it may be noted that it is an argument in favour of 30725 being directly the source, that displacements by small multiples of the oun have fitted in so remarkably closely and frequently in the formation of the list of lines above.

Taking then the limits as $30725{\cdot}30 + \xi$, $32589{\cdot}40 + \xi + d\nu_1$, $33419{\cdot}04 + \xi + d\nu_1 + d\nu_2$ the denominators are found to be

$$1{\cdot}889322 - 30{\cdot}74\xi$$

$$54831 - 2{\cdot}60\xi + 28{\cdot}14d\nu_1$$

$$1{\cdot}834491 - 28{\cdot}14(\xi + d\nu_1)$$

$$22914 - 1{\cdot}14\xi - 1{\cdot}1d\nu_1 + 27d\nu_2$$

$$1{\cdot}811577 - 27(\xi + d\nu_1) - 27d\nu_2$$

In these ξ cannot be greater than a few decimals and will produce no effect on the differences of the mantissæ, which may themselves be affected with errors ±1 due to using 7-figure logs.

The differences may be represented as follows :—

$$5\,(10996{\cdot}8+5{\cdot}6d\nu_1-{\cdot}52\xi\pm{\cdot}2)-\delta_1 = 5\Delta_2-\delta_1$$

$$2\,(10999{\cdot}0-{\cdot}5d\nu_1+13{\cdot}5d\nu_2\pm{\cdot}5)+6\delta_1 = 2\Delta_2+6\delta_1,$$

where to a first approximation we know already Δ_2 is 10998 ± 1. The sum

$$= 7\,\{10997{\cdot}4-{\cdot}53\xi\pm{\cdot}1+4\,(d\nu_1+d\nu_2)\}+5\delta_1.$$

Quite small changes in $d\nu_1$, $d\nu_2$ can therefore make the connections with Δ_2 exact, and since the multiples 5, 2, 7 are so small any small errors in the approximate value of Δ_2 can have no effect. If we use the ν_1 as found from the occurrency curve 1864·5 $d\nu_1 = {\cdot}4$, the number in the first bracket is 10999.

The agreement with both is so close to the relation indicated that it speaks strongly in support of the D origin of the limit, and the outstanding small differences may be left for the present.

If now the other supposition be tested, viz., that the separations arise from a S (∞) source $= 51025{\cdot}29+\xi$, the three denominators are found to be

$$\begin{array}{ll} 1{\cdot}466091-14{\cdot}36\xi & \\ & 26067-{\cdot}75\xi+13{\cdot}61d\nu_1 \\ 1{\cdot}440024-13{\cdot}61\,(\xi+d\nu_1) & \\ & 11160-{\cdot}31\xi-{\cdot}31d\nu_1+13{\cdot}30d\nu_2 \\ 1{\cdot}428884-13{\cdot}30\,(\xi+d\nu_1)-13{\cdot}30d\nu_2, & \end{array}$$

and in the 20312 set the S (∞) must enter as a negative quantity since the separations are there in inverse order. In this case ξ may be considerable. The differences may be expressed as follows :—

$$2\Delta_2+6\tfrac{3}{4}\delta-53-{\cdot}75\xi+13{\cdot}61d\nu_1\,;\;\; \Delta_2+\delta_1+9-{\cdot}31\xi+13{\cdot}3d\nu_2.$$

No permissible values of ν_1, ν_2 can make these both multiples of the oun. If it is possible to do so by a proper choice of ξ the latter must satisfy $53+{\cdot}75\xi = 153m$, $-9+{\cdot}31\xi = 153n$, where m, n are integers and 153 is the value of the oun. This requires $22 = 47{\cdot}4m-115n$. A suitable solution is $m = -2$, $n = -1$, which requires $\xi = 480$, or, say, series limit $= (11\tfrac{1}{4}\delta)$ S (∞). This method of explanation looks then improbable especially when taken with the more natural one above. It may be concluded with some confidence that the series in question is of the F type depending on D series for limits as in the usual way.

In the suggested lines for $m = 1$, found by sounding with $2e$ the mantissa of $f(1)$ was found to be $90\,\{10998{\cdot}8-{\cdot}4\xi+{\cdot}4dn\}-\delta$. ξ is small and the term involving it may be omitted. The error dn in 3010 may, however, amount to a few units because the lines on which it was based were assumed to depend on the limit 30725, whereas there is the possibility that they might belong to one of the parallel series found in the F sets depending on $(x\delta_1)$ (30725). In case, however, of dn being small and the mantissa involving the term in δ, we might expect still to find lines depending on the $90\Delta_2$, as the presence of the δ suggests satellites. To get $90\Delta_2$, requires a $+\delta$ displacement which decreases $f(1)$ by 17·02. In other words lines with wave-numbers 17·02 larger for F and less for **F**. We do not find this completely, but the following sets are observed, already given in the notes to the list under $m = 1$.

$2e.F_1$	$2e.F_2$	$2e.F_3$
(6) 17638·55	(3) 19503·29	(1) 20333·22
	12·52	
....	(3) 19515·81	
25·53		
(5) 17664·07		

$3e.\mathbf{F}_1$	$3e.\mathbf{F}_2$	$3e.\mathbf{F}_3$
[36498·90]	[38363·40]	(1) 39193·40
5·08		
(1) 36493·82		
	17·64	**17·66**
....	(2) 38345·76	(2) 39175·74
25·45	**26·31**	
(3) 36473·45	(1) 38337.09	

In which permissible observation errors are $dn = \pm{\cdot}7$. As has been seen the 36493 corresponds to a δ_1 displacement in the limit. The others to $3\delta_1$, $6\delta_1$, and δ in the sequent $f(1)$. The lines 38345, 39175 consequently have their sequent mantissa exactly $90\Delta_2$.

Further it was found that $(3n)$ 39683·95 is $2e.v.(-\delta_1)$ $\mathbf{F}_1(1)$. The next preceding line to this is (1) 39666·49 or 17·46 behind it, again showing the required δ displacement and having the $90\Delta_2$ mantissa.

If it be granted that the series is of the F type, the limit must be a d-sequent. Consequently the mantissa of $30725{\cdot}30+\xi$ must be a multiple of the oun. Its mantissa is $889322-30{\cdot}74\xi = 81\,(10998{\cdot}13-{\cdot}38\xi)-10\delta_1 = 81\Delta_2-10\delta_1$ with great exactness. Let the true value of Δ_2 be $10998{\cdot}20+x$. Then if the relation is exact $81x+30{\cdot}74\xi+5{\cdot}7 = 0$ or $\xi = -2{\cdot}63x-{\cdot}18$, $x = -{\cdot}38\xi-{\cdot}07$. Now we know that ξ must be a small fraction, certainly $<\pm{\cdot}5$. Hence x must lie between $\pm{\cdot}2$ and $\Delta_2 = 10998{\cdot}20\pm{\cdot}20$. We should, therefore, expect this value for Δ_2 except possibly where electronic changes of atomic weight came in, as has been suggested above. If then the 1864 separations depend on exact $5\Delta_2-\delta_1$ and $2\Delta_2+6\delta_1$ we get as closer approximations $5{\cdot}6d\nu_1 = 1{\cdot}40$ or $d\nu_1 = {\cdot}25$ and ${\cdot}80+13{\cdot}5d\nu_2 = 0$ or $d\nu_2 = -{\cdot}05$, in other words $\nu_1 = 1864{\cdot}35$, $\nu_2 = 829{\cdot}59$ when the limit is 30725. When this limit is displaced by $y\delta_1$ these change by $\cdot45y$, $\cdot22y$.

Further, the conditions for $f(1)$ require $\cdot67 - x - \cdot4\xi + \cdot4dn = 0$, or $dn = -1\cdot8 - \cdot056\xi$. If the sounder $2e$ was exactly normal this dn must be due to observation errors in 17638 of $d\lambda = \cdot5$, but the value of e is also subject to some small uncertainty. In any case the result shows that the reference line does not depend on a displaced 30725, for if so dn would be at least 4·97.

Returning now to the discussion of the D series let us consider first this second group of clearly analogous series of lines :—

1.	(1) 19942·53	**1775·45**	(10) 21717·98	4.	(10) 20636·30	**1767·09**	(20) 22403·39	**813·66**→(<1)
2.	(1) 19989·72	**1780·27**	(6*n*) 21769·99	5.	(2) 20688·96	**1785·54**	(10) 22474·44	
3.	(1) 20017·46	**1783·71**	(10) 21801·17	6.	(1) 20859·23	**1784·57**	(7) 22643·80	
	(1) 20021·66	**1779·51**		7.	(4) 20962·07	**1780·09**	(10) 22742·16	**813·80**→(1)

They all, with the doubtful exception of 4, 7 have the appearance of belonging to first, or doublet, satellite sets, in which the second line is always the stronger. The 1780 separations are clearly associated with the now well recognised mid-triplet abnormality. That it is not itself a normal separation is indicated by No. 3 in which the 1779·51 also occurs.

In (1) the separation 1775·45 is $\nu_1 - 2\cdot45$. It differs from the displaced $(\delta)\,\nu_1$ by ·29 which is within error limits. In this case the limit would be $(\delta)\,\mathrm{D}(\infty)$ which is 42·48 less than $\mathrm{D}(\infty)$ and $= 50982\cdot81$. With this limit the mantissa of 19942 comes to $879711 = 80 \times 10998\cdot3 - \delta_1$ or $80\Delta_2 - \delta_1$ within error limits. This is the typical form for the second satellite set of a triplet D series, but modified by the δ_1 displacement, so common in this group of elements, though here it appears in an apparently first satellite set instead of the second. We note at present that taking account of the small corrections, and writing as before $\Delta_2 = 10998\cdot2 + x$ its true value is $80\Delta_2 - \delta_1 + 8 - 30\cdot28\xi + 30dn - 80x$. The observation error dn is $< \cdot2$ in this region and ξ is probably < 1.

In (2) 19989 is 47·19 above (1). The change due to the displacement δ in the limit is 42·55, whilst δ_1 in the sequent gives 5·05 suggesting that the limit of (2) is the normal $\mathrm{D}(\infty)$, with sequent $80\Delta_2$. With this limit the mantissa is found to be $879853 = 80 \times 10998\cdot16$, or with small corrections $80\Delta_2 - 3\cdot2 - 30\cdot29\xi + 30dn - 80x$.

In (3) we have the modified 1783·71 with the clearly real separation 1779·51 or $\nu_1 + 1\cdot61$. Now the displacement due to -3δ on the ν is 1·60 which is practically exact. This gives a limit 31·88 larger or 51057·17, and the mantissa becomes 879855, the same as for (2) and $= 80\Delta_2 - 1\cdot4 - 30\cdot29\xi + 30dn - 80x$. The line 20017·46 is 4·20 behind the other. A displacement of δ_1 in the sequent term produces 5·05. Thus 20017 is very close to a line with mantissa $= 80\Delta_2 - \delta_1$, but the difference ·85, corresponding to $d\lambda = \cdot21$, is too great to render the relation exact.

Nos. 4 to 7, although much further towards the violet should not be put aside. They all show the exceptional separations.

In (4) the separation is 1767·09 = $\nu_1-10·81$. If it corresponds to a real ν_1 the limit will be (5δ) D (∞) which reduces ν_1 by 10·7 and makes the limit about 212 less. The mantissa would be $82\Delta_2+7\frac{1}{2}\delta+71$. This difference (71) from an exact multiple of δ_1 shows this to be impossible.

The discussion of the 1864 series has definitely shown it to be of the F type, and has given the limit within very small errors. This limit is one of the *d* (1) sequents of the diffuse series. Its mantissa was found to be $81\Delta_2-2\frac{1}{2}\delta$ or $80\Delta_2+15\frac{1}{2}\delta$. It gives one firm starting point for the discovery of the D series. The results just obtained indicate the lines which through their dependence on multiples of Δ_2 give the origin term of the diffuse sequence. They point, as we have seen, to the existence of several displaced, or parallel, sets of diffuse series. It is possible to show definitely that these exist, even if there be some uncertainty as to the lines occupying the position of D_{11} (1). In the normal case with a single diffuse series, the D_{11} (1) is always the strongest line of the series. Also in the normal type the D limit is the same as that of S—here $51025·29+\xi$. When however displacement occurs, the energy of a single line is dispersed amongst several others, and a line corresponding to the normal may be weak, or even too faint to have been observed. As a matter of experience also it is found that the lines of low order (m = 1, 2...) are subject to these displacements in a much greater degree than those for higher orders of *m*. Now there are a number of lines, which by their position and absence of ν_1 separations to stronger lines have the appearance of being D_{11} (1) lines. If they are, their mantissæ must differ from multiples of Δ_2 (in the present case $80\Delta_2$) by multiples of the oun. The fact that they may do so does not of course prove that they are D_{11} lines. If they do not do so it proves that they are not. They may however in the latter case belong to a displaced series, satisfying the multiple law when the proper displaced limit $(y\delta_1)$ S (∞) is employed. This gives us a method of testing as to what displacement a given line may correspond. If our calculus were already fully established the next step would be to apply this test to the above lines. But in reality we are testing our calculus to see if it can be firmly established, and our immediate aim must be to obtain independent evidence for the existence of parallel series. For this immediate purpose it will only be necessary to apply the test to two lines, the general question being postponed for the present.

In the first attempt at arranging the D_{11} series the strong lines (8) 20559·08 and (10) 38366·36 were taken for m = 1, 2, and the formula calculated with the limit D (∞) = S (∞). As will be seen immediately, this gave satisfactory agreement with sounded observed lines up to m = 15; and as a matter of fact this series was used to test for the parallel sets displaced $(\pm 2\delta_1)$ S (∞) on either side of it. Now the formula constants for a set only vary slightly if the wave-number of the line chosen for m = 1 is changed considerably. This therefore did not prove definitely that 20559 is the correct D_{11} (1), and as a fact it does not satisfy the multiple test. Its mantissa is $897337-31·14\xi$. That of the line 19989, which is shown above to be the origin of the normal D set is $879853-30·29\xi$. The difference is 17484 = $28\frac{1}{2}\delta+71$. This is as far as it can be from being a true multiple, and it is quite impossible to explain this by any observation errors in the two lines. For instance our maximum admitted error can only change the mantissa by 6. If the test is valid therefore 20559 cannot be $D_{11}(1)$. Yet it has all the appearance of such a line. Is it a displaced one? Suppose it corresponds to $y\delta_1$. Now a displacement of δ_1 on the limit changes its value by 10·62, so that $(y\delta_1)D(\infty) = 51025{\cdot}29+\xi-10{\cdot}26y$. If p denote the ratio of actual observation error to the maximum permissible, *i.e.*, $O = d\lambda = {\cdot}05p$, the mantissa of 20559 with the new limit is $897337+330{\cdot}7y+6p-31{\cdot}14\xi$. The denominator of 19989 is $80\Delta_2 = 879856+80x$. Also $330{\cdot}7 = \frac{1}{2}\delta+25{\cdot}7$. Hence the mantissa of 20559 is

$$\begin{aligned}&80\Delta_2-80x+17481+2y\delta_1+25{\cdot}7y+6p-31{\cdot}14\xi\\&\quad= 81\Delta_2+10\tfrac{1}{2}\delta+2y\delta_1+68-81x+25{\cdot}7y+6p-31{\cdot}14\xi\end{aligned}$$

in which x is small (about $\pm{\cdot}2$). Also from the consideration of 19989 above

$$-3{\cdot}2-80x+6p'-30{\cdot}29\xi = 0.$$

Eliminating x

$$71+25{\cdot}7y+6(p-p')-{\cdot}47\xi = M(153)$$

in which $p, p'<1$ and ξ cannot exceed about 2.

It is not possible to satisfy this with $y = 0, \pm1$, or ±2.

With

$$\begin{aligned}y &= 3 \qquad -6+6(p-p')-{\cdot}47\xi = 0,\\ y &= -3 \qquad -5+6(p-p')-{\cdot}47\xi = 0.\end{aligned}$$

If, then, 20559 be a D_{11} line it belongs to one of $(\pm3\delta_1)D(\infty)$, and is definitely excluded as a possible normal D_{11}.

The next line of higher frequency is the weak line (1) 20581·64. Its mantissa is $898040-31{\cdot}17\xi = 81\Delta_2+11\frac{3}{4}\delta+7-81x+6{\cdot}5p-31{\cdot}17\xi = 81\Delta_2-11\frac{3}{4}\delta$ within error limits. This therefore passes the D_{11} test. If it is the actual D_{11} its weak intensity is due to the numerous displacements for $m = 1$. If it be taken as $D_{11}(1)$ with the previously mentioned (10) 38366·36, and the limit $D(\infty)$, the series formula is found to be

$$n = 51025{\cdot}29-N\Big/\left\{m-{\cdot}988854-\frac{{\cdot}090814}{m}\right\}^2.$$

The lines after $m = 2$ lie in the violet outside the observed region. To test them therefore recourse must be had to sounding, only the *e.u.v.* links have been used for this purpose. The results are given in the middle column of the subjoined table and exhibited in diagram (Plate 3). Details are given in the notes following the table. Lines were calculated down to $m = 15$ and tested. The result may be regarded as conclusive in establishing the series, as well as increasing confidence in the method of sounding—a confidence which reposes not on a single coincidence, but on the recurrence of a large number of successive ones. As will be seen the agreement between the calculated values and those found by sounding is remarkably close. The only doubtful case may be that for $m = 1$. If 20559 be taken for this the formula is only slightly changed and the agreement almost as good, the calculated value for $m = 3$ giving O−C $d\lambda$ = ·05 instead of ·00. It is excluded because it cannot belong to a diffuse series in which the limit is S (∞).

TABLE of Parallel XD Series.

After $m = 2$ the wave-numbers at the head of each set are those calculated from the formula and are not, as usual, enclosed in [ ]. The links are entered as attached to the respective observed lines. On the right of each is entered the O−C ($d\lambda$) calculated as if the error were on the observed line. They would be less—often considerably less if calculated on the series line itself.

m.	$(2\delta_1)$ D (∞).		D (∞).		$(-2\delta_1)$ D (∞).	
1	[20550·58], = $(-2\delta_1)$	**31·06**	(1) 20581·64	**31·06**	[20610·70], = $(+2\delta_1)$	
	e.(2) 27864·16	·06			*e*.(6) 27926·53	− ·22
	u.(6) 24683·37	·06			*u*.(<1) 24749·35 $(-\delta_1)$	− ·11
	v.(2) 24977·40	− ·18			*v*.(5) 25036·99	·26
2	(2) 38345·76*	**20·60**	(10) 38366·36	**21·20**	[38387·61]	
	(1) 31033·15 (δ_1).*e*	− ·02			(<1) 34251·50.*u*	·02
	(5) 33915·20 $(-2\delta_1)$.*v*	·00			(2) 33062·43.*v*	− ·02
	or				*v*.(5) 35499·89.*e*	·12
	[38345·11]	**21·25**			(26940·63).*e*.*u*	− ·05
	(<1) 34212·13.*u*	− ·01			*or* [38390·24] = $(2\delta_1)$	
	(4) 33908·99 $(6\delta_1)$.*v*	− ·09			(<1) 31076·51.*e*	− ·07
3	44004·50 = $-\delta_1$	**21·94**	44026·48	**21·76**	44048·20	
	(<1) 36690·52.*e*	− ·01	(1) 36712·34.*e*	0	*u*.(1) 40867·04.*e*	·01
	(1) 32262·65.*e*.*v*	− ·02	(2) 39598·95.*v*	− ·02	(2) 32304·23 $(-2\delta_1)$.*e*.*v*	·03
4	46557·99	**20·24**	46578·23	**21·43**	46599·66	
	(3) 34815·79.*e*.*v*	·00	[39264·13.*e*]		(1) 31971·46.2*e*	·00
			(<1) 34836·78.*e*.*v*	− ·05	(1) 27836·93.2*e*.*u*	·17
			(35130·95).*e*.*u*	0		
5	47927·39	**21·39**	47948·78	**21·45**	47970·23	
	(<1) 28871·19.2*e*.*v*	·00	(<1) 33322·45.2*e*	− ·18	(2) 36523·41.*e*.*u*	− ·03
	(3) 33301·47.2*e* ?	− ·20	(3) 36209·04.*e*.*v*	− ·24	(3) 36229·78.*e*.*v*	− ·13
			(1) 29189·92.2*e*.*u*	− ·29		
6	48748·78	**21·35**	48770·13	**21·35**	48791·48	
	(1) 37300·91.*e*.*u*	·04	(<1) 34140·88.2*e*	·09	(4) 34163·39.2*e*	·09
	(3*n*) 29692·79.2*e*.*v*	− ·02	(1) 29710·88.2*e*.*v*′	− ·06	*v*.(3) 38589·92.2*e*	·09
			(2) 37320·68.*e*.*u*	·15		
			(5) 30005·34.2*e*.*u*′	− ·01		
7	49280·00	**21·25**	49301·25	**21·25**	49322·50	
	(3) 37829·35.*e* $(-\delta_1)$.*u*	·00	(1) 34870·17.2*e*	·24	*u*.(1) 30559·83.2*e*	·14
	(2) 37534·58.*e* $(-\delta_1)$.*v*	·00	(3) 37849·82.*e*.*u*	·29		
	v.(1) 39082·95.2*e*	− ·21	(<1) 37554·87.*e*.*v*	·32		
			(<1*n*) 30444·37.2*e*.*v*	·07		

---

* This is also 3*e*.$\mathbf{F}_2$ (1) (δ) the 90 Δ_2 linked $\mathbf{F}_2$. It is probably not $(2\delta_1)$ D (2) but hides it.

TABLE of Parallel XD Series (continued).

m.	$(2\delta_1)$ D(∞).		D(∞).		$(-2\delta_1)$ D(∞).	
8	49643·23	**21·25**	49664·48	**21·25**	49685·73	
	(3) 37901·76.*e.v*	−·04	(<1) 35037·31.2*e*	−·08	*u*.(1) 39193·40.2*e* ?	−·19
	(<1) 35011·06		(1) 38217·68.*e.u*	−·03		
	15·47.2*e*	−·03				
	(4*n*) 35019·89		(1) 30607·90.2*e.v*	·04		
	(1) 30879·26					
	81·78.2*e.u*	·00				
	(1) 30884·31					
	(2) 30588·24.2*e.v*	−·13				
9	49902·63	**21·25**	49923·88	**21·25**	49945·31	
	(1) 38159·20.*e.v*	·09	(35295·68).2*e*	00	(1) 38203·51.*e.v*	−·03
	(3) 31139·94.2*e.u*	·13			*v*.(1) 39745·63.2*e*	−·04
10	50094·30	**21·25**	50115·55	**21·25**	50136·80	
	u.(2) 39598·95.2*e*	·02			(1) 38689·81.*e.u*	−·03
	(2) 31035·76.2*e.v*	·24	(1) 38666·62.*e.u*	·11	(<1*n*) 31076·84.2*e* $(-\delta_1)$.*v*	·07
	or *e.e* (δ_1).*v*	·0	(31355·28).2*e.u*			
11	50239·95	**21·25**	50261·20	**21·25**	50282·45	
	(3) 38791·71.*e.u*	·06	(1) 35637·28.2*e* (δ_1)	·03	(1*n*) 38835·86.*e.u*	−·04
	(4) 35610·00.2*e*	·14	(1) 38521·55.*e.v*	·16	*v*.(3) 40082·56.2*e*	·00
	u.(1) 39745·63.2*e*	−·04				
	(3) 31479·21.2*e.u*	−·07				
	(<1) 31180·63.*e.e* $(-\delta_1)$.*v*	·09			(3) 31224·82.*e.e* $(-\delta_1)$.*v*	·01
12	50353·24	**21·25**	50374·49	**21·25**	50395·74	
	u.(1) 38906·95.*e*	−·06	(2) 35745·69.2*e*	·05	(5) 35767·81.2*e*	−·02
	(1) 31592·28.2*e.u*	−·04	(1) 38632·72.*e.v*	−·02	(1) 31636·86.*e.e* (δ_1).*u*	−·02
13	50443·04	**21·25**	50464·29	**21·25**	50485·54	
	(3) 31680·16.2*e.u.*	·15	(1) 35836·52.2*e*	−·03	(<1) 38935·20	
					38·00.*e.u.*	·01
	(5) 31384·86.*e.e* $(-\delta_1)$.*v*	−·03	(1*n*) 38720·67.*e.v*	·10	(1) 38940·81	
			u.(2) 39969·62.2*e*	·02	(6) 31725·89.2*e.u* (δ_1)	·00
					(<1) 31431·51.2*e.v* (δ_1)	−·03
14	50515·50	**21·25**	50536·75	**21·25**	50558·00	
	(3) 38970·66.*e* $(-\delta_1)$.*u*	·02	(3) 31479·21.2*e.v*	·13	(1) 39107·56.*e* $(-\delta_1)$.*v*	·05
	(1) 35889·37.2*e* ?	−·16			(1) 35929·09.2*e*	·06
	(1*n*) 31755·50.2*e.u*	−·13			*u*.(1) 40063·94.2*e*	−·06
15	50574·68	**21·25**	50595·93	**21·25**	50617·18	
	(1*n*) 38835·86.*e* (δ_1).*v*	−·06	(2) 35963·72.2*e*	·30	(3) 35985·33.2*e* $(-\delta_1)$	−·07
	(1) 35949·76.*e* (δ_1).*e*	−·07	*or* 2*e* $(-\delta_1)$	−·05	(1*n*) 31755·50.2*e.u*	·03
	u.(3) 40082·56.*e* (δ_1).*e*	−·02	(1) 38852·76.*e.v*	·06	(<1) 31559·98.2*e.v*	·10
	(5) 31717·13.*e* (δ_1).*e.u*	−·05				

Notes on Table of D(∞).—*m* = 3. The *e* linked line has separation 1780·29 to (4*n*) 38492. *m* = 4. The *e* linked line is not observed, but it would be separated 1781·55 from (1*n*) 41045. 35130. There is no line to this, but it seems split into two as indicated in the scheme

(2) 35126·05					(1) 37720·05	
	31·05	**1776·95**	(3) 36908·00	**816·17**		24·17
(2) 35136·05					(1) 37728·30	

In which it may be noticed that the sum of the separations is 2593·12, the normal value. The order m is too large to definitely settle the sequence displacements if in them. They would be close to $\pm 18\delta_1$ for the first and $\pm 14\delta_1$ for the third. But if the connection is real a better explanation might be modification of the links, $\pm\delta_1$ on $e.u$ give the exact numerical agreement for the two D_{33} lines, and $e(\pm\delta_1)$, $u(\pm 2\delta_1)$ would give 5·86 where 6 is observed.

$m = 5$. 33322 has separations **1782·14** (4), **811·06** (4), the sum being normal.

$m = 6$. 34140 has separations 1782·92 to (2) 35923. Further, there is 30005 **1780·45** (1) 31785 **813·4** (1) 32599, in which the sum of the separations is closely normal. The $2e.u$ and $2e.v$ linked lines differ by 3·41 and 3·05 from the calculated, but are correct if the links $u' = u(-2\delta_1)$ and $v' = v(-2\delta_1)$ are taken. They are inserted since the corresponding normal links occur in other lines and the amounts are exact.

$m = 7$. The $2e$ linked 34870 has triplet separations **1784·44** (3), **813·73** (3), also the $2e.v$ has **1778·14** to intensity (<1). The $e.u$, $e.v$ linked lines are inserted although their difference as they stand are so considerable, because they all are exact if the links involved are taken as displaced $(-\delta_1)$.

$m = 9$. There is no observed $2e$ linked line, but it seems to have split up into two, thus

(<1) 25293·16 **1778·20** (4) 37071·36
95·65
(<1) 25298·14

The two observed lines are numerically $D_{11}(9)(\pm\Delta_2)$.

$m = 10$. The $2e.u$ line is split up into two (<1) 31350·91, (<1) 31359·66, of which the former shows **1778·79** (4), **811·75** (5).

It should be noticed how many of the e and $2e$ linked lines introduce the modified triplet separations.

Notes on Table of $(\pm 2\delta_1)D$.—$m = 1$. The displacements on the limits would give 20560·40 and 20602·89. There is the already noted 20559 near the first discarded for $D(\infty)$ because it does not pass the multiple test. It serves better for $(-2\delta_1)D$, but would require at least an observation error $d\lambda = -\cdot 1$ which we have regarded as excessive. There are no observed lines connected by e, u, v links to either, nor near them. Those given in the table are, however, very clear. They make the sequence term displaced $2\delta_1$ from that of the D series, viz., $-2\delta_1$ for $+2\delta_1$ on limit, and $+2\delta_1$ for $-2\delta_1$ on limit, *i.e.*, interchange of $\pm 2\delta_1$ for $\mp 2\delta_1$ on limit. In the third series the e and v linked lines differ respectively by $e + 1\cdot 74$ and $v - 1\cdot 70$. They form therefore a parallel inequality, and are good evidence in spite of the considerable difference 1·7.

$m = 2$. The limit separation should give for the first series 38345·11, and it apparently exists although possibly it belongs to a series commencing with 20559. There is clear evidence of a set corresponding to $m = 1$, shown by sounding and giving a sequence displacement of $-6\delta_1$. In the third series the line depending on the limit change alone would be 38387·56. This gives links $u + 2\cdot 85$ and $v - 2\cdot 87$ with the lines indicated or a parallel inequality.. They are explained by $\pm 2\delta_1$ displacements in the sequent.

$m = 3$. Here δ_1 as a displacement in the sequent produces a separation of ·5. The sequent displacement in the first series is therefore $-\delta_1$ and $+\delta_1$ in the third. The line 32304 however shows $-2\delta_1$.

$m = 7$. Modified links $e(-\delta_1)$ are introduced. This is supported by the fact that the two lines given differ respectively by 2·37, 2·32 from values given by normal e, whilst the modification of e by δ_1 produces 2·32. The double example and exact difference give weight to the suggestion.

$m = 10, 11$. Similarly the modified e makes exact agreement, and they enter in a corresponding way and in both series.

We can now use this series to test the question as to the existence of parallel series depending on $(\pm 2\delta_1)$ D (∞). This does not mean that the sequences must be the same in each. In fact it is to be expected that there may be a concomitant change in them also, but they can only differ by a few multiples of the oun. The important point to notice is that for large values of m, the effects of such differences become negligible and the observed separations from the standard D should approximate to $\pm 21 \cdot 25$, which is that due to $2\delta_1$ on the limit. It will not be necessary to go into such a full discussion as in the standard series with D (∞), except for the first three sets, where the evidence for changed sequence terms with displaced limits is conclusive and important. The lists and sounders are given in the same table as for the standard series, in the first and third columns respectively. The considerations adduced enable us to feel that the ground is safe in recognising that parallel series depending on displacements on the limit D (∞) = S (∞) really exist. The evidence does not depend on a single numerical coincidence of a line, or of a line found by sounding, but on the fact that these numerical coincidences appear for so many sounding links in all the 15 sets tested. This is not affected by the high probability that several of the sounded lines are chance coincidences. This result then gives more confidence to the method partially applied above, in the application of the law that the D sequences must have mantissæ which differ by multiples of the oun from multiples of Δ_2, or in other words must be themselves multiples of the oun. This method consisted in testing certain lines to see whether by using the displaced limits $(y\delta_1)$ D (∞)—now seen to really exist—the above relation holds.

The evidence seems to show that the typical lines—D (∞) = S (∞)—for $m = 1$ have been much affected by displacement effects, and that consequently the intensities of the normal lines themselves are much diminished. Although this is some disadvantage, it will be well to attempt here to get some insight into the complete satellite system for the first two orders.

We have seen that 19989 belongs to this normal set with a mantissa = $80\Delta_2$ and that 20581 satisfies the condition necessary for a D_1 line with this. The difference of their mantissa (see below) is $29\frac{1}{4}\delta$. If they are of the D_{12}, D_{11} types, as is indicated by the fact that the first belongs to a doublet and the second stands by itself, a triplet satellite set should be expected whose first line D_{13} is separated from the D_{12} by about three-fifths that of D_{12} from D_{11}. Its mantissa should therefore be about $18\delta = \Delta_2$ less. This would mean a line about 19623· forming the first line of a triplet. No line is observed here. There are, however, lines at (1) 19602·66 and (3) 19632·44 of which 19602 passes the suitability test for a normal D line, and the other does not. The mantissa of 19602 is $79\Delta_2 - \delta$, *i.e.*, 19δ behind that of 19989 and rather too large. On the other hand the problematic 19623 may be too weak, in which case the corresponding D_2, D_3 lines which should be stronger might be observable. The D_2 line should be about 21400. We do find this, in fact, with triplets of a kind. The whole set of these lines can then be arranged as follows :—

(1) 19602·66	**1774·45**	(3n) 21377·11	**821·04**	(2) 22198·15
δ				
[19623·05]	**1777·90**	(2) 21400·95	**809·53**	(2) 22210·48
18δ				
(1) 19989·72	**1780·27**	(6n) 21769·99		
29¾δ				
(1) 20581·64				

in which it may be noticed that in the first the sum of the separations is the same as 1780·19+815·30, *i.e.*, the modified $\nu_1+\nu_2$, and in the second 809·53 = 815·20−5·67 whilst a δ_1 displacement on the sequent produces a change of 5·14. With the limit 51025·29+ξ the mantissæ of the *d* sequents are—$d\lambda = {\cdot}05p$—

$$\begin{aligned}
19602,\quad & 868240-29{\cdot}73\xi+6p = 79\times10998{\cdot}2-618+6p-29{\cdot}73\xi \\
19623,\quad & 868846-29{\cdot}75\xi+6p' = 79\times10998{\cdot}2-12+6p'-29{\cdot}75\xi \\
19989,\quad & 879853-30{\cdot}29\xi+6p = 80\times10998{\cdot}2-3+6p-30{\cdot}29\xi \\
20581,\quad & 898040-31{\cdot}17\xi+6{\cdot}5p = 80\times10998{\cdot}2+18184+6{\cdot}5p-31{\cdot}17\xi
\end{aligned}$$

in which p lies between ±1 and p' depends on error of extrapolated line, and may be >1. Writing as before $10998{\cdot}2 = \Delta_2-x$ these become respectively

$$\begin{aligned}
&79\Delta_2-\delta-7+6p-79x-29{\cdot}73\xi \\
&79\Delta_2\ -12+6p'-79x-29{\cdot}75\xi \\
&80\Delta_2\ \ -3+6p-80x-30{\cdot}29\xi \\
&80\Delta_2+29\tfrac{3}{4}\delta+7+6{\cdot}5p-80x-31{\cdot}17\xi
\end{aligned}$$

The multiple rule requires that the last four terms in each expression must vanish. This is clearly possible for small values of p, say $<\frac{1}{2}$, and a single relation between x and ξ, say $8x = -3\xi$.

As a further test of the reality of the extrapolated line 19623 linked lines may be sought for. There is none for $+e$, but lines are found close for u, v, $e\pm v$, viz.,

(6) 23754·27	**1778·40**	(3) with $u-2{\cdot}1$	19623·19
($<1n$) 24053·37		„ $v+2{\cdot}1$	19623·16
(2) 24509·92		„ $e-v$	19623·82
(2) 31362·71		„ $e+v$	19620·66

Taking 19623 as D_{13}, the satellite differences are 29¾δ and 18δ. Since 18 × 5 = 90 and 29¾ × 3 = 89·25, these separations are very closely in the ratio 5 : 3 in accordance with the rule for the known triplet series in other groups. The triplet set 19602 appears somewhat anomalous. The middle line appears to have the modification so common in the middle set of a triplet, but in the opposite direction to the usual one, *i.e.*, the sequent to the second line is $d(-\delta_1)$ instead of $d(\delta_1)$.

The preceding considerations have shown the existence of a complete set of D (1) satellites. But further, the D suitability test shows that 20305·60 may be a D_1 line with normal limit and that with it may go two sets of extrapolated triplets in which the mantissa of the first is a multiple of Δ_2. They are

[16013·45]	**1777·90**	(1) 17791·35	**816·44**	(8) 18607·79
	1784·83	(1) 17798·28	**809·51**	
[17725·39]	**1777·90**	(3) 19503·29	**816·96**	(4) 20320·25
(1) 20305·60				

The first set, however, involves for the D_{33} the line 18607, which belongs to the 1864 F set discussed above, and is stronger than we should expect. Provisionally we will suppose the true D_{33} is hidden by the F line. The line 20320 has been already adduced (p. 383) as showing connection with the 1864 separations in a similar way to 20312. The mantissæ of the D_1 lines are (p' indicating extrapolated lines), $769890 = 70 \times 10998{\cdot}2 + 16 + 3{\cdot}2p' - 25{\cdot}2\xi = 70\Delta_2 + 16 + 3{\cdot}2p' - 70x - 25{\cdot}2\xi$, $814815 + 4{\cdot}2p' - 27{\cdot}25\xi$, and $889493 + 6p - 30{\cdot}75\xi$. They differ successively by $44925 = 4 \times 10998{\cdot}2 + 1\frac{1}{2}\delta + 16 + 4{\cdot}2\,(p'_2 - p'_1) - 2{\cdot}0\xi = 73\frac{1}{2}\delta + 16 + 4{\cdot}2\,(p'_2 - p'_1 - 4x - 2{\cdot}0\xi$ and $74679 = 122\frac{1}{4}\delta - 17 + 6p - 4{\cdot}2p'_2 - 7x - 3{\cdot}5\xi$. The conditions are satisfied within error limits that the mantissa of the extreme satellite is $70\Delta_2$, and that the differences for the satellites are due to $122\frac{1}{4}\delta$ and $73\frac{1}{2}\delta$. Since $122\frac{1}{4} \times 3 = 366{\cdot}7$ and $73\frac{1}{2} \times 5 = 367{\cdot}2$, the normal ratio of satellite separations is again reproduced. The evidence is clear, therefore, for two groups of normal diffuse series depending on $79\Delta_2$ and $70\Delta_2$ respectively.

The fact that 20320 is connected with 1864 in the same way as 20312, that the ν_2 separation is not good, and that we should expect a doublet here in place of a triplet rather point to the supposition that it is not a member of the set. For the D_{13} line the sequent is $51025{\cdot}29 - 16013{\cdot}45 = 35011{\cdot}84$, on this the oun displacement produces a change of 6·04 which accounts for the modified $\nu_1 = 1784{\cdot}83$ in the usual way. The mantissæ of the two D_{11} groups 20581, 20320 differ by 14δ.

It will be sufficient here to attempt the allocation of the corresponding satellites for $m = 2$ only. They should be at about the same oun multiple distance from the D_{11} line as for $m = 1$. The $D_{11}(2)$ has been taken as 38366 with denominator 1·943447. Taking the first group, the satellites should have denominators about $29\frac{3}{4}\delta$ and $47\frac{3}{4}\delta$ less than this. These would correspond to 38208·61 and 38111·69. With regard to the first the lines in this neighbourhood are (1) 38199·58, (1) 38203·51, (1) 38217·68, of which the first and third are respectively 9·03 less and 9·07 greater than 38208, and suggest equal displacements on either side. They are found to correspond to $31\frac{1}{2}\delta$, 28δ from D_{11}, or $\pm 7\delta_1$ on either side of 38208, with errors

$d\lambda = \pm{\cdot}02$, whilst 38203 corresponds to $30\frac{3}{4}\delta$ with $d\lambda = {\cdot}01$. They should form portions of doublets with lines at 39977·48, 39995·58. None have been observed, but it must be remembered that these are close to the end of the observed region where only the stronger lines would be seen. Sounding with $-(e+u)$ we find lines (1) 28530·91, (3) 28548·26 respectively, $e+u-{\cdot}71$ and $e+u+{\cdot}04$ behind the expected D_{22} lines showing no lines at ν_2 ahead, which should be visible if they existed. We may, therefore, conclude the two lines in question belong to a doublet set. With regard to the suspected D_{13} line (38111) we find in this neighbourhood

(1) 38134·61	**1780·23**	[39915·84]	**815·20**	(1*n*) 40731·04
(1) 38129·37	**1778·21**	(1) 39907·58	**810·36**	(1) 40717·94
(1*n*) 38108·60				

Calculation shows that the first lines of these sets give with $D_{11}(2)$ separations depending on displacements $43\frac{1}{2}\delta\,(d\lambda = 0)$, $44\frac{1}{2}\delta\,(d\lambda = {\cdot}01)$, $48\frac{1}{4}\delta\,(d\lambda = -{\cdot}03)$. All that can be said is that these lines may be the required satellite sets. As, however, a displacement of $\frac{1}{2}\delta_1$ only produces a change of about ·05 in λ it is not possible to get any certainty. It is further possible that some of the set may belong to parallel series. For instance $38108 = (2\delta_1)\ 38129$. The first two are, however, so close $d\lambda = {\cdot}00$ and ·01 that they are entered as D_{13} lines.

For the second group 20305 is a 14δ displacement from 20581. We should expect the corresponding $m = 2$ line about the same displacement from 38366. This is satisfied by the line (1) 38292·32 giving 14δ with $d\lambda = -{\cdot}01$. There is also a doublet (1) 38285·86, **1778·08**, (1) 40063·94 with normal separation and displaced $15\frac{1}{4}\delta$ with $d\lambda = {\cdot}00$. The first D (1) is displaced $122\frac{1}{2}\delta$ from its D_{11} line 20305 and the second by an extra $73\frac{1}{2}\delta$. If the $D_{11}(2)$ line is taken as 38285 the calculated line with $122\frac{1}{2}\delta$ is 37611·59. There is no line here, but there is a doublet, which with this line may be written

(2) 37606·15	**1780·83**	(1) 39386·98
[37611·59]	**1775·39**	

The mean of these two separations is 1778·11 or the normal ν_1. We have here a parallel inequality due to $2\delta_1$ displacement in 39386 and an extra $2\delta_1$ in 37606, or 37606 is $123\frac{1}{2}\delta$ ahead of D_{11}.

Again calculating the $73\frac{1}{2}\delta$ displacement on 30706 the line should be 37174·40. It is not observed but there is a triplet close to it showing a similar inequality to the former. It is

(1) 37159·39	**1774·38**	(1) 38940·81	**813·67**	(<1) 39754·48
(1*n*) 37166·43	**1781·42**			

The mean of the two first separations is 1777·90 or exact ν_1 and we have an exact parallel inequality. The inequality is due to two successive $1\frac{1}{4}\delta$ displacements, and 37159 is exactly $2\frac{1}{2}\delta$ extra on the calculated 37174, or 76δ from the first satellite.

The first two orders of the two groups are represented in the following scheme in which the satellite separations are given as sequent displacements from d_{11}:—

The $79\Delta_2$ group.				The $70\Delta_2$ group.		
			$m = 1$.			
[19623·05]	(2) 21400·95	(2) 22210·48		[16013·45]	(1) 17791·35	(8) 18607
$\mathbf{47\frac{3}{4}}\delta$		$\mathbf{48}\delta$		$\mathbf{195\frac{3}{4}}\delta$		
(1) 19989·72	($6n$) 21769·99			[17725·39]	(3) 19503·29	
$\mathbf{29\frac{3}{4}}\delta$				$\mathbf{122\frac{1}{4}}\delta$		
(1) 20581·61			$\mathbf{14}\delta$	(1) 20305·60		
			$m = 2$.			
(1) 38129·37	(1) 39907·58	(1) 40717·94		(1) 37159·39		
$\mathbf{44\frac{1}{2}}\delta$	$\mathbf{44\frac{1}{2}}\delta$	$\mathbf{46\frac{1}{4}}\delta$		$\mathbf{199\frac{1}{2}}\delta$	} (1) 38940·81	(<1) 39754·48
				($1n$) 37166·43		
				$\mathbf{197}\delta$	$\mathbf{198\frac{1}{4}}\delta$	$\mathbf{198\frac{3}{4}}\delta$
(1) 38199·58						
$\mathbf{31\frac{1}{2}}\delta$				(2) 37606·15	(1) 39386·98	
(1) 38217·68	(3) 28548·26.*e.u*			$\mathbf{123\frac{1}{2}}\delta$	$\mathbf{122\frac{1}{2}}\delta$	
$\mathbf{28}\delta$	$\mathbf{28}\delta$					
(10) 38366·36			$\mathbf{14}\delta$	(1) 38285·86		

Without dealing with the whole of the material at disposal we will illustrate its application by considering in more detail the portion of the spectrum given on p. 382 in which the majority of the lines undoubtedly belong to D (1) systems. It is to be noticed that the effectiveness of the method in the present case depends on the facts, (1) that the observation errors do not exceed $d\lambda = ·05$, and (2) that with $m = 1$ it is consequently possible to determine the values of the mantissæ to within 6 units in the sixth significant figures, whilst a displacement of one oun in the sequent produces a change in λ of the order 1·2, or twenty-four times the maximum observation error. The limit 51025 being supposed displaced by $y\delta_1$ becomes $51025·29-10·62y+\xi$. The mantissæ of the sequences are then calculated with this limit, and expressed in terms of Δ_2, δ_1, x and p where $\Delta_2 = 10998·2+x$, and $-p$ is the ratio of the observation error to the maximum ($d\lambda = ·05$). The series more fully discussed above is definitely taken as depending on the limit $y = 0$. In other words the mantissa of 19989 is exactly $80\Delta_2$ which condition requires, writing q for its p,

$$3·2+30·29\xi-6q+80x = 0,$$

and gives a relation between ξ and x. The term in x in each mantissa is then eliminated by means of it. There can be little doubt about the allocation of 19889, but even should it be in error the doubt does not affect the argument as to the relative displacements of the different lines, as $y\delta_1$ denotes the displacements relative to 19889. The final results are given in the following table :—

1	19880	$80\Delta_2+2y\delta_1-5\frac{1}{2}\delta+66+14y+{\cdot}2\xi+6(q-p)$	**–4,**	11	**–5,**	–4
2	19942	$80\Delta_2-\delta_1$ with $y=4$	**4,**	11		
3	19959	$80\Delta_2+2y\delta_1-\delta-61+16y+{\cdot}05\xi+$ „	**4,**	3		
4	19989	$80\Delta_2$ with $y=0$	**0**			
5	20017	$80\Delta_2+2y\delta_1+1\frac{1}{4}\delta+77+17y-{\cdot}05\xi+$ „	**–4,**	9	**4,**	–8
6	20021	„ $+1\frac{1}{2}\delta+52+17y-{\cdot}05\xi+$ „	**–3,**	1	**6**	
7	20029	„ $+2\delta-31+17y-{\cdot}05\xi+$ „	**2,**	3		
8	20041	„ $+2\frac{1}{2}\delta+50+17y-{\cdot}07\xi+$ „	**–3,**	–1	**6,**	0
9	20080	„ $+6\frac{1}{2}\delta-2+18y-{\cdot}11\xi+$ „	**0,**	–2		
10	20107	„ $+10\delta-33+18y-{\cdot}17\xi+$ „	**2,**	3		
11	20305	„ $+15\frac{3}{4}\delta+14+21y-{\cdot}46\xi+$ „	**0,**	14	**–1,**	–7
12	20312	„ $+16\delta+82+21y-{\cdot}47\xi+$ „	**–4,**	–2	**3,**	–8
13	20320	„ $+16\frac{1}{2}\delta+9+21y-{\cdot}47\xi+$ „	**0,**	9		
14	20333	„ $+17\frac{1}{4}\delta-50+22y-{\cdot}50\xi+$ „	**2,**	6	**–5,**	–7
15	20443	$81\Delta_2+2y\delta_1+4\frac{3}{4}\delta-11+24y-{\cdot}30\xi+$ „	**0,**	–11	**–6,**	–2
16	20454	„ $+5\frac{1}{4}\delta+41+24y-{\cdot}30\xi+$ „	**–2,**	–7	**5,**	8
17	20467	„ $+6\delta-12+24y-{\cdot}45\xi+$ „	**0,**	–12	**–6,**	–3
18	20470	„ $+6\delta+76+24y-{\cdot}45\xi+$ „	**–3,**	4	**3,**	–5
19	20500	„ $+7\frac{1}{2}\delta+71+26y-{\cdot}48\xi+$ „	**–3,**	–7	**3,**	–4
20	20529	„ $+9\delta+80+26y-{\cdot}48\xi+$ „	**–3,**	2	**3,**	5
21	20559	„ $+10\frac{1}{2}\delta+71+26y-{\cdot}48\xi+$ „	**–3,**	–7	**3,**	–4
22	20581	„ $+11\frac{3}{4}\delta+10+26y-{\cdot}51\xi+$ „	**0,**	10		
23	20596	„ $+12\frac{1}{2}\delta+11+26y+{\cdot}53\xi+$ „	**0,**	11		
24	20636	„ $+14\frac{1}{2}\delta+36+26y-{\cdot}62\xi+$ „	**–1,**	13	**4,**	–13

Of the above (14, 18) must be set aside at once: (14) because it belongs to the triplet set linked by $2e$ to the parallel set F (1), and (18) because it is F_2 (2). It may also be noted that neither have the prevalent separation 1780 to lines of higher frequency. Of the numbers on the right of the list, those in thick type give the values of y which bring the outstanding differences to the corresponding number in ordinary type. These differences must be due to errors either in ξ or observation. Since ξ is small, it is seen that they must be capable of annulment by the observation errors $6(q-p)$, and must, therefore, at the maximum be <12. The smallness of ξ can be seen from the following considerations which connect it with the corresponding ξ (say ξ') for the 1864 F series. The limit of the F is $30725{\cdot}30+\xi'$ with mantissa $889322-30{\cdot}74\xi'$. It is a $d(1)$ sequent. If 19989 is a D line with $y=0$, its sequent is $51025{\cdot}29+\xi'-19989{\cdot}72=32035{\cdot}57+\xi'$. The mantissa of this is $879853-30{\cdot}29\xi'+6q$. Both being d sequents must differ by a multiple of the oun. Their difference is $9469+30{\cdot}29\xi-30{\cdot}74\xi'-6q$ and $15\frac{1}{4}\delta=9470$. Hence

$$30{\cdot}29\xi-30{\cdot}74\xi'=1+6q,$$

or,

$$\xi=1{\cdot}015\xi'+{\cdot}03+{\cdot}2q.$$

Thus $\xi=\xi'\pm{\cdot}3$. Since the 1864 limit is determined as the mean of F and **F** series its value is subject to a very small uncertainty and ξ' will be a small fraction.

Consequently in the above list the value of ξ has no importance in settling the order of the displacements.

The whole of the foregoing argument is based on the constancy of Δ_2 for all series. This matter has been referred to in the introduction in which the question of what is to be understood by the atomic weight was brought up. The accuracy in the determination of the oun is rendered so great by the constitution of the d and f sequences, that the mass of the electrons connected with the nucleus affect it. It may be interesting to illustrate the considerations there adduced by a concrete example. The example we will take is (1), as the result may possibly throw some light on a difficulty which will appear later. The line 19880 is seen to require a displacement of $-\delta$. Let us determine the transfer of electrons in order to produce a change by one oun. Suppose this transfer changes x to $x+x'$. This means that the mantissa as represented in the list must be diminished by $80x'$, by putting $y = -1$.

Hence

$$x' = \frac{52}{80} = {\cdot}65, \qquad d\delta = {\cdot}036.$$

The change in the number of electrons (see p. 342) = 925 × ·036/611 × 130 = 1·97 = 2. The addition of eight electrons to the mass acting in our standard case, would render 19980 a possible D line with limit $(-\delta_1)\,\mathrm{D}(\infty)$ instead of a possible one with $(-\delta)\,\mathrm{D}(\infty)$.

The preceding treatment of the material is only a first step towards unravelling the intricacies of these D systems. An exhaustive treatment is here impossible, and would involve the consideration of other data—the triplet separations, linkages, similarities of arrangement, dependence on F series and so on. All that can be done is to give a few illustrative instances and to bring into prominence certain problems whose solution in the future may be of extreme importance.

(1) The line 20021 is given as requiring the displacement $-3\delta_1$ in the limit, and in this case the mantissa is $80\Delta_2$. This displacement increases the limit 51025 by 31·88, and the resulting ν_1 should be greater by 3 × ·535 = 1·60 and = 1779·50. As a fact the line (see p. 382) forms a doublet with separation 1779·51, and with intensities 1, 10, in the proper order for a D_{12} set. All the tests support each other. Again in the doublets

(7) 20500·13	**1784·68**	(2) 22284·82
(8) 20559·08	**1785·64**	(<1*n*) 22344·72

the same displacements (or of $+3\delta_1$) are indicated. If we suppose these modified values of ν_1 produced in the same way as in Kr by a relative displacement in the sequent, the oun in this case alters the separation by 4·91. The modified ν_1 therefore becomes 1779·50 + 4·91 = 1784·41, or within error limits of the observed value for 20500, but too small for 20559. That for 20559 corresponds to an extra displacement of $-2\delta_1$ on the limit, or $-5\delta_1$ in all, which is quite inadmissible on the qualifying test.

The mantissæ of both lines differ by an exact 3δ. This is the natural conclusion, viz., same limit, sequents differ by 3δ, but then the value of 1785 remains unexplained, and we should not expect to find two D_{11} lines of the same group so close together. A possible explanation is to allot 20559 to the alternative displacement of $3\delta_1$ which gives $\nu_1 = 1776\cdot30$ and to take $2\delta_1$ in the sequent of the second line. This would give a modified separation $1776\cdot30+2\times4\cdot91 = 1786\cdot12$ or ·48 greater than that observed. It may be noted that as they stand the four lines (18 to 21 of list) requiring $-3\delta_1$ displacement have their sequence mantissæ equally spaced by $1\frac{1}{2}\delta$. The line 20041 also comes into the system with an exact δ sequence displacement from 20021.

In an analogous position to the 16013 line in the $D(\infty)$ set appears the $(-3\delta_1)D(\infty)$ set

[16044·05] **1785·37** (1) 17829·41 **810·13** (2) 18639·54

in which the 16044 is extrapolated from 18639 by the $\nu_1+\nu_2 = 2595\cdot5$. As the oun displacement on the sequent here produces 6·06, the modified ν_1 should be $1779\cdot50+6\cdot06 = 1785\cdot56$, and the modified $\nu_2 = 816\cdot02-6\cdot06 = 809\cdot96$ which agree with the observed. We may regard the 16044 as subject to the observation errors of 18639, say $d\lambda$. The mantissa of 16044 is $769863 = 70\ (10998\cdot04-\cdot36d\lambda)$, or $70\Delta_2$ as in the case of $D(\infty)$.

With the examples here given the arguments from the capability test and the observed separation agree. Complete sets for $(-3\delta_1)\ D(\infty)$ series are obtainable but are not here adduced.

(2) The line of 19942 has already been considered on (p. 396) as a $(4\delta_1)\ D(\infty)$. It again is a case where the capability test and ν_1 separation both point to the limit displacement of $4\delta_1$ or δ. The list also gives 19959 as requiring the same displacement. But it is $u.F_4(3)$ of the series below and shows a forward link 1864 in analogy with these F series.

(3) The line 20312 is clearly of special importance. It gives the source of the 1864 separations. As is seen the capability test reqires $-4\delta_1$ but there are many difficulties in the way of properly placing it. It forms part of a strong doublet (6) 20312 **1783·72** (5) 22096 in which the intensities are not in normal order. If 1783·72 is the usual modified ν_1, the sequence oun displacement is here 4·98 and the true $\nu_1 = 1783\cdot72-4\cdot178 = 1778\cdot74$. This differs only by ·13 from that proper to $(-2\delta_1)\ D(\infty)$, instead of to the $-4\delta_1$. The 1864 separations are greater than in the F series, being 1865·51, 832·13 against those found from the F above 1865·10, 829·64. The latter were based on the limit 30725·30 on which a displacement of δ_1 produces ·45 in 1864 and ·22 in 829. There is clearly here some triplet modification as the ratio of the separations is not correct. With 1865·51 should go 830·15 or with the given sum 1866·87, 830·77. Now $-3\delta_1$ on the limit produces 1865·45 and increase the limit itself to 30740·21. This should be the d sequent for 20312. If so the D limit should

be 30740·21+20312·70 = 51052·91 which is 27·62 above the normal and has no reference to the oun displacement. If 20312·70 corresponds to the $-4\delta_1$ indicated, the limit is 51067·79 and the *d* sequent is the difference, or 30755·09. This is the source of the 1864 separations. It is 22·79 greater than the limit 30725·30 of the F series considered in connection with this series and which gave 1864·10 and 829·64. This limit corresponds to a sequent $6\delta_1$ less displacement or 29·82 greater value which is practically exact. This present sequent will therefore increase 1864·10 by 6 × ·45 to 1866·70 and 829·64 by 6 × ·22 to 830·96. The new sum is 2697·66, as against the observed 2697·44 in remarkably close agreement. This then supports the $-4\delta_1$ indication of the capability test. A $-3\delta_1$ displacement would make the sum = 2697·00 —which though a worse agreement may yet be within observation errors. The same $-3\delta_1$ would make the modified $\nu_1 = 1779{\cdot}50+4{\cdot}97 = 1784{\cdot}47$, or ·72 greater than the observed. One is almost tempted to suspect here an error greater than the ordinary observation error. An error of ·36, $d\lambda = -{\cdot}09$, would make all three tests agree in allotting 20312 to the $-3\delta_1$ set, and would bring it into the group of lines considered in Case 1.

It should be noted that this 20312 line is the line with wave-length 4922 referred to by LIVEING and DEWAR for its peculiar behaviour (see p. 350).

(4) The separations of the triplet 19880 are 1778·42, 815·30. They suggest the separation $-\delta_1$ to which belong 1778·43, 815·42. The capability test gives $-3\delta_1$ and the two are incompatible. In the D series given below the calculated limit from the first three lines give the limit as 51045·37 or 20·08 above the normal. With the uncertainty in a limit found in this way this is a displacement of $-2\delta_1$ which gives 21·25. Further there seems some evidence to show that this line with 20021, 20041, 20312, 20500, 20559 belong to one D (1) group of lines. The evidence consists in the existence of parallel F sets showing displacements equal to the separations of these lines. In other words these lines are D sets with the same limit. All these tests mutually exclude each other. How can their indications be reconciled? I suggest

(1) The limit for the line is $(-\delta_1)$ D(∞) and the observed ν_1, ν_2 are thus explained.
(2) The capability test is met by the transfer of six electrons indicated above.
(3) It is not the first line of the series in question.
(4) The F separations will be found to offer a natural explanation.

In what has preceded an attempt has been made to allocate normal D series, but they are clearly not the strongest sets. In my first attack on the D, F problem the procedure adopted was to take the 19880 triplet as a clear satellite set, and attempt by the application of RYDBERG's table to find a series for this satellite series. The three sets found were :—

(1) 19880	**1778·42**	(5) 21659·14	**815·30**	(10) 22474·44
[37888·59]	**1777·90**	(1) 39666·49	**814·42**	(2) 40480·91
[43801·13]	**1777·74**	[45578·87]	**818·18**	[46397·05]

The second set is near the end of the observed region, and the fainter D_1 line is extrapolated by the normal ν_1, this will introduce a small possible error. The third set is wholly outside the observed region and was obtained by sounding. The data are as follows, the lines used being regarded as linked by *e* to the first triplet 36487, ...

[43801·13]	$d\lambda$	[45578·87]	$d\lambda$	[46397·05]	
(1) 36487·03.*e*	·00	(1*n*) 38264·77.*e*	·00	(1) 39082·95.*e*	·00
(1) 29175·18*e.e* (δ_1)	·00	(1) 30949·88.2*e*	·08	(<1) 31768·93.2*e*	−·01
(1) 39372·71.*v*	−·02	(1) 41155·14 *v* ($2\delta_1$)	·03	(<1) 34659·40.*e.v* ($2\delta_1$)	−·07
(1) 39666·49.*u*	·09	(1) 33841·74.*e.v* ($2\delta_1$)*	·12		
(3) 32351·67.*e.u* ($-\delta_1$)	·10	(1) 34126·78.*e.u* ($-2\delta_1$)*	−·12		
(<1) 24749·35.*e.e* (δ_1).*v* (δ_1)	−·04	(1) 26814·18.2*e.u* ($2\delta_1$)	·03		

The formula obtained from these is

$$n = 51045{\cdot}37 - \mathrm{N}\Big/\left\{m + {\cdot}898460 - \frac{{\cdot}022504}{m}\right\}^2.$$

This was tested to $m = 13$, with good agreement with the exception of $m = 4$. All are in the ultra-violet and require sounding. As the evidence of the efficacy of sounding already adduced may be regarded as convincing, there is no advantage in giving further details of the results, especially as no additional conclusions are based on this formula.

XF.—We are to look for parallel F sets with the same separations as those of the D satellites. Starting with the triplet 19880 as an undoubted D satellite triplet it was then attempted to determine the others, by picking out those lines of larger wave-length whose mantissæ differed from that of 19880 by oun multiples. The data are contained implicitly in the list on p. 407, independently of the form into which the mantissæ are there thrown. It is clear that the above condition is satisfied by all those which show unsatisfied remainders of the same magnitude as that of 19880—*i.e.*, 66—within error limits of, say, ±6. The selected lines were 19880, 20021, 20041, 20312, 20500, 20559. We have shown above that the second, third, fifth and sixth of these satisfy the capability test for $-3\delta_1$ displacements on the limits, and have also found some evidence that 20312 also belongs to this set, although perhaps with a somewhat excessive observation error. It has also appeared that 19880 cannot belong to this set, but it has so happened that the satellite separations from it reproduce themselves in the F series sufficiently well as to serve for identification. We shall, therefore, use this allocation. The selected lines give separations from 19880 respectively of 140·94, 161·01, 431·98, 619·41, 679·36. In these 19880 was treated as if it were a fifth satellite to 20559 and the notation adopted of F_6, ... F_1 series. Such series were found, and in what follows I shall chiefly confine consideration to them leaving aside for the present purpose, with one or two exceptions, the very numerous other groups which exist. It would not be advisable to suggest a definite notation for these groups until the whole system of D and F series is placed on a secure and comprehensive foundation. The above notation, therefore, is only to be regarded as one which in this communication serves to identify certain of those groups here more specially discussed.

---

* Parallel inequality.

The material was treated by the same method as in Kr, and search made for the above separations in the region where F (2) lines should be expected. From the lines thus obtained the actual F lines were sought for, bearing in mind the large sequence displacements so common in the low orders. In only three cases—those for $F_{2,3,6}$—were suitable undisplaced lines found for $m = 2$, and of these only F_3, F_6 gave observed lines for $m = 3$. They were—separations from F_6—

m.	F_6.		F_3.		F_2.
2.	(8) 18812·55	**432·46**	(8) 18380·09	**619·82**	(1) 18192·73
3.	(1) 24253·30	**432·49**	(<1) 23820·81		
5.	(2) 28108·52	**432·00**	(1) 27676·52		

For $m = 4$, F_3 showed no undisplaced line, but in good position for $m = 5$ there were lines for F_6, F_3. From the three lines for F_3 the following formula was found

$$30740·17 + \xi - N \Big/ \left\{ m + ·986181 - 622\xi - \frac{·014730 - 1003\xi}{m} \right\}^2 .$$

We can at once apply two tests to this as to fulfilling the conditions for F series. The limit 30740 must be a *d*-sequent, *i.e.*, must differ by oun multiple displacement from some other known *d*-sequent, and the first *f*-sequent must depend directly on a multiple of Δ_2. We already have a very accurate *d*-sequent found as the limit of the 1864 series, viz., $30725·30 + \xi'$. On this a δ_1 displacement produces a change of 4·97 so that $-3\delta_1$ produces $30725·30 + \xi' + 3 \times 4·97 = 30740·21 + \xi'$, and this condition is accurately satisfied by $\xi_1 = \xi' + ·04$ where ξ_1 belongs to the present case. For the second test the mantissa of $f(2)$ is

$$978816 - 121\xi_1 = 89\,(10997·93 - 1·36\xi_1) = 89\,(\Delta_2 - ·27 - ·36\xi_1 - x)$$

which again is seen to satisfy the test exactly. The satisfaction of these two conditions must give full confidence as to the correctness of the allocation of the F_3 lines.

From this formula lines were calculated from $m = 4$ to 27 and the other sets allocated by their corresponding separations. They are not reproduced here, however, beyond $m = 15$, partly because they are only of importance in a systematic arrangement of the X spectrum, and partly because the lines to be identified are so close that it becomes a matter of extreme difficulty to allocate them correctly. For instance, the calculated values of F_2 (14) and F_1 (15) are 30064·38 and 64·16; of F_4(12) and F_3(16) are 30360·47 and 60·29 with many other examples. With the high orders successive lines become close, and with the large number of separate series involved the observed spectrum should be expected to be crowded, as indeed in this region it is.

The results are given in the annexed table, with notes up to $m = 7$. Numbers in brackets refer to values obtained—when none have been directly observed—by connection with linkages or displacement. It may be recalled that a link from an observed line to an expected but unseen one leads to the inference that the unobserved really exists, whereas a displaced line, when the displacement is on the limit, gives evidence only for the value of the sequent $f(m)$, and when the displacement is on the f-sequent, is evidence to that effect alone.

The true $F_3(\infty)$ is 30740·17+ξ where ξ is small. The other limits will depend on oun displacements from this. Estimated from 20312—the source of $F_3(\infty)$—these displacements, expressed as multiples of δ_1 are $-21\frac{1}{2}$, $-14\frac{1}{2}$, $-13\frac{1}{2}$, $9\frac{1}{2}$, $12\frac{1}{2}$. Their values can, therefore, be calculated with exactness relatively to $F_3(\infty)$. They are

F_1.	F_2.	F_3.	F_4.	F_5.	F_6.
493·11 + 1·012ξ	30552·13 + 1·01ξ	30740·17 + ξ	31010·41 + ·99ξ	31030·58 + ·98ξ	31172·24 + ·98ξ
4·910	4·927	4·964	5·030	5·043	5·090
	59·02	247·06	517·30	537·47	679·13

The numbers below the limits give respectively the changes produced in them by the displacement of one oun. The numbers in the last line give the calculated accurate separations of the corresponding F lines from F_1.

For the first order, $m = 2$, considerable displacements are to be expected. Only normal lines for $F_{2,3,6}$ are observed. The set (5) 18998·40, (3) 19515·81, (7) 19676·25 give close normal separations 517·41, 677·85. Now the limit of F_1 is 30493·11 and the denominator of 18998 calculated from this is 3·088906−134·3ξ, or a mantissa = 1088906−134·3ξ+24p = 99 (10998·14−1·35ξ+·24p) = 99Δ_2. The normal $f(2)$ sequent is 89Δ_2. There is, therefore, a displacement of 10Δ_2 in the sequence term. Further the defect in the separation 677·85 from the normal is 1·28 whilst a δ_1 displacement on the sequent produces 1·11. The lines in question, therefore, are $F_1(2)(10\Delta_2)$, $F_4(2)(10\Delta_2)$, $F_6(2)(10\Delta_2-\delta_1)$.

To find a representative for F_5 we may test (1) 18332·41, **290·35**, (3) 18622·76 and (1) 18018·31, **289·60**, (6) 18307·91 in which the separations refer to that of F_3, F_5, viz., 290·40. The mantissæ difference for the first set is 5740, and the nearest oun multiple $9\frac{1}{2}\delta = 5804$ is outside error limits. That of the second is 42758 and $4\Delta_2-2\delta = 70\delta = 42770$. An observation error of $d\lambda$ = ·03 would make this exact. The lines in question may therefore be $F_3(2)(70\delta)$, $F_5(2)(70\delta)$. So also it may be shown that 18466·47 is $F_2(2)(-2\Delta_2)$.

With the F difference-series occur also the **F** summation type. As their existence is a new fact of great importance the evidence available up to $m = 10$ is given. The results are embodied with those of F in the table.

TABLE of the $F_n.F_n$ Lines

	n = 1.	ν_2.	n = 2.	ν_3.	n = 3.
Limit	30493·11	**59·02**	30552·13	**247·06**	30740·17
m. 2	[18133·42] **93·63** (42853·85) (1*n*) 2581.*u* (5) 5262 18998·40 }	**59·31** **57·68** F $(10\Delta_2)$	(1) 5495 18192·73 **52·13** (42911·53) (δ_1) (1) 2808.*e*	**246·67** **247·91**	(8) 5439 18380·09 **40·92** (43101·76) ? (< 1) 3187.*e.v*
3	(3) 4240 23576·11 **92·72** 37409·33 (3) 2672·35	**59·21** **59·01**	*v*.(3) 3562 (23635·32) **51·83** 37468·34 (3) 2668·14	**244·70** **251·29**	(< 1) 4196 23820·81 **40·74** (37660·62) (− 3δ_1) (1) 2655·57
4	(2*n*) 4617.*v* (26078·04) **92·12** 34906·21 (< 1) 2864	**59·25** **57·98**	(− 2δ_1) (1) 3826 (26137·29) **50·74** (34964·19) (3δ_1) (1) 2858	**247·86** **247·92**	*v*.(3) 3250 (26325·90) **40·01** (35154·13) (− 2δ_1) (3) 2844
5	(2) 3644 27430·33 **91·83** 33553·33 (6) 2979	**58·49** **60·86**	(− δ_1) (2) 3636 (27488·82) **51·50** (33614·19) (δ_1) (3) 2973	**246·19** **251·18**	(1) 3612 27676·52 **40·51** (33804·51) (− δ_1) (5) 2957
6	(1) 4197.*v* (28242·73) **94·09** (32745·46) (1) 3530·40.*v*	**58·23** **59·26**	(δ_1) 3531·93 (28300·96) **52·49** (32804·72) (− δ_1) (1*n*) 3047	**247·06** **246·90**	(1) 3509 28489·79 **41·07** (32992·35) (− δ) (1*n*) 3031
7	(1) 3474 28773·67 **93·06** (32212·45)	**58·57** **59·96**	(5) 3467 28832·24 **52·33** (32272·41)† (− 2δ_1) (1) 3098	**245·65** **247·29**	(1) 3445 29019·32 **39·53** 32459·74 (4) 3079

---

* $F_6(7)$ *and* $F_2(9)$ *are coincident.*

from $n = 1$ to $n = 6$.

ν_3.	$n = 4$.	ν_5.	$n = 5$.	ν_6.	$n = 6$.
517·30	31010·41	**537·47**	31030·58	**679·13**	31172·24
					(8) 5314·15
				679·13	18812·55
					72·34
517·86	(43371·71)	**536·27**	(43390·12)	**678·29**	(43532·14)
	(4) 2772.*e*		(1*n*) 2546.*u*		(δ_1) (< 1) 2759.*e*
	(3) 5122 }				(7) 5080 }
517·41	19515·81 }			**677·85**	19676·25 }
	(1) 5008.*u*		(5) 4145		(1) 4122
516·71	(24092·82)	**537·71**	24113·82	**677·19**	24253·30
	10·58		**31·52**		**71·22**
519·02	37928·35	**538·89**	37949·22	**679·81**	38089·14
	(1) 2635		(3) 2634		(1) 2624
	u.(5) 3253		(1) 3756		(1) 3736
517·35	(26595·39)	**537·59**	26615·63	**680·03**	26758·07
	08·22				**72·37**
515·84	35421·05			**680·47**	35586·68
	(1) 2822				(3) 2809
	(5) 3576		(1) 3574		(2) 3556
519·81	27950·14	**537·32**	27967·65	**678·19**	28108·52
	10·50		**30·57**		**72·33**
517·53	(34070·86)	**540·17**	34093·50	**682·82**	34236·15
	($3\delta_1$) (4) 2932·72		(3) 2932·27		(3) 2920
	(δ_1) (1) 3475		(< 1*n*) 4056.*u*		(< 1) 4033.*u*
517·55	(28760·28)	**537·14**	(28779·87)	**678·89**	(28921·62)
	10·33		**30·41**		**71·78**
514·93	(33260·39)	**535·50**	33380·96	**676·36**	(33421·95)
	($2\delta_1$) (1) 3004·81		(2) 2994		(δ_1) (1) 2990·74
	(1*n*) 3413		($2\delta_1$) (< 1*n*) 3409		($-\delta_1$) (2) 3394
514·91	29288·58	**536·94**	(29310·61)	**678·95**	(29452·62)*
	10·79		**30·34**		**72·48**
520·56	(32733·01)	**537·63**	(32750·08)	**679·90**	(32892·35)
	($-\delta_1$) (4) 3054		($-2\delta_1$) (2) 3494.*u*		(δ) (1) 3037

† Coincident values.

TABLE of the $F_n.F_n$ Lines

	$n = 1.$	$\nu_2.$	$n = 2.$	$\nu_3.$	$n = 3.$
$m.$ 8	(1n) 3998.u 29134·56 92·83 (31851·09) (−δ_1)(1) 3139‡	58·74 59·46	(1) 4569.e 29193·30 51·92 31910·55 (1) 3132	246·56 247·11	e.(2) 3070.u 29381·12 39·66 32098·20 (4) 3114
9	(2δ_1)(1) 3400 (29393·35) 92·82 31592·28 (1) 3164·43	59·15 59·01	(−δ_1)(2) 3394 (29452·50)* 51·90 (31651·29) (δ_1)(4) 3632.u	247·58 248·87	e.(δ_1) 3045.u (29640·93) 41·04 (31841·15) δ_1(1) 3139‡
10	(2) 3379 29584·48 92·58 (31400·68) (−2δ_1)(3) 3184·74	60·21 59·04	(3) 3918.u (29644·69) 52·20 (31459·72) (δ_1)(2) 3177	248·12 247·61	(δ_1)(1n) 3350 § (29832·60) 40·44 (31648·29) (−2δ_1)(2) 3160
11	(1) 3362 29727·58	59·41	(1) 3942.v * (29786·96)	247·35	(δ)(5) 3332** (29974·93)
12	(3) 3349 29843·17	58·88	v.(5) 2912 (29902·05)	246·94	v.(<1) 2896 (30090·11)
13 F(2Δ_2) or	(1) 3339 29934·65 (δ_1)(1) 3339·37 29932·42	60·16 59·57	(5) 3332 29994·81 (−2δ_1)(1) 3334 29991·99	247·01 247·34	(<1) 3312 30181·66 (−2δ_1)(1n) 3313 30179·76
14	(5) 3331 30005·34		[30064·38]	247·17	(2δ_1)(2) 3303 30252·51
15	[30064·16]	59·04	(1) 3318†† 30123·20	248·14	(<1) 3298 30312·30

* $F_6(7)$ *and* $F_2(9)$ *are coincident.*

† Coincident values.

‡ May be either.

§ *Coincident with* (−2δ_1) $F_1(12)$ *of* 1864 *series.*

from $n = 1$ to $n = 6$ (continued).

ν_3.	$n = 4$.	ν_5.	$n = 5$.	ν_6.	$n = 6$.
517·37 516·92	[29651·93] 09·97 (32368·01) $(-\delta_1)$ (1) 3089	536·65 537·09	[29671·21] 29·70 (32388·18) (δ_1) (3) 3537.*u*	677·80 689·72	$(-2\delta_1)$ (1) 3354 (29812·36) 71·58 (32530·81) $(-\delta_1)$ (4) 3073
516·71 518·30	*v*.(<1) 2911·38 (29910·06) 10·32 (32110·58) (δ_1) (3) 3112	537·23 538·45	$(2\delta_1)$ (2) 3339·00 (29930·58) 30·65 (32130·73) $(-3\delta_1)$ (3) 3112	679·52 680·55	*u*.(<1) 2922·62 (30072·87) 72·87 (32272·83)† $(-2\delta_1)$ (1) 3098
516·70 519·93	$(2\delta_1)$ (6) 3322§ (30101·18) 10·90 (31920·61) $(-2\delta_1)$ (1) 3132¶	538·72 537·00	(1) 3318‖ 30123·20 30·44 (31937·68) (2) 3595.*u*	678·20 680·16	(2) 3303‖ 30262·68 71·76 (32080·84) (δ_1) (1) 3614.*v*
516·61	$(-\delta_1)$ (4) 3306 (30244·19)	535·10	(2) 3303†† 30262·68	677·17	(5) 3288 30404·75
517·73	*v*.(2) 2873 (30360·90)	539·32	(<1) 3290 30382·49	678·26	$(-\delta_1)$ (3) 3276 (30521·43)
516·79 516·98	$(3\delta_1)$ (1) 3281 (30451·44) $(-\delta_1)$ (<1*n*) 3283 30449·40	538·38 538·01	(4) 3280·66 30473·03 (<1) 3280·94 30470·43	678·54 678·89	*v*.(δ_1) (2*n*) 2852 30613·19 $(2\delta_1)$ (4) 3264 30611·31
	[30522·68]	538·53	(3) 3273 30543·87	679·30	(1) 3258 30684·64
517·33	$(-\delta_1)$ (<1) 3269 (30581·49)	537·17	(δ) (4) 3264 (30601·33)	677·56	$(2\delta_1)$ (3) 3250 30743·72

‖ F_6 (10) and F_5 (11) are coincident.

¶ This line is F_2 (8).

** 3332 is F_2 (13).

†† F_5 (10) and F_2 (15) coincident.

In this table under each order the first line gives the wave-length of the observed line to the last Ångstrom, its intensity, and, where necessary, the displacement or linkage to be applied. The second line gives the wave-numbers of the F lines and the thick type the separations from the F_1 line adopted. Up to $m = 10$ the fourth line gives in the same way the wave-numbers of the F lines and the fifth the corresponding wave-lengths. In the third line the numbers give the mean limit $\frac{1}{2}(F+F)$, but only the last four significant figures are entered, the complete calculated values being given at the head of the table.

Notes to Table.—$m = 2$. For F_3 in addition to that given there are (-2δ) (1) 38632·72.v = 43101·20, $(3\delta_1)$ (1) 38687·71.v = 01·53, (δ) (1) 38988·33.u = 01·27, (δ_1) (<1) 38973·59.u = 01·71.

$m = 3$. The linked F_2 agrees with $(3\delta_1)$ (2) 23650·79 = ...35·79. The linked F_4 with $(3\delta_1)$ (1) 24077·41 and $(-6\delta_1)$ (1) 24062·12 both of which give the same value ...92·30.

$m = 4$. The linked F_1 agrees with $(-6\delta_1)$ (1) 26048·86 = ...78·32. For F_2, $(2\delta_1)$ (1) 26147·89 = ...38·06 is closer to the calculated value ...37·95. F_4 has a link v = 4428·62. For F_6, $(-2\delta_1)$ (3) 26744·54 = ...54·72 is only ·26 greater than the calculated value. Most of the observed lines of this order are one or two units larger than the calculated. F_3 is also given by $(3\delta_1)$ (<1) 37675·75 = 60·66. For F_4, $(-\delta_1)$ (2) 35417·16 = ...22·19 gives separation correct. For F_6 $(-\delta_1)$ (1) 35580·98 = ...86·07 gives much closer separation.

$m = 5$. For F_2, $(2\delta_1)$ 27498·89 = ...88·03. For F_4, e.35261·18 = 27947·18 and $(3\delta_1)$ 27963·59 = ...48·50 both give better separations. F_5 shows a series inequality with $-(u+3·28)$ and $u-2·93$. For F_6 also (δ_1) 28113·58 = ...08·49 and (δ) 28133·98 = ...13·62. F_1 is a strong observed line which makes the mean limit ...91·83 too small and some of the other separations too large. $(-2\delta_1)$ 33544·55 = ...54·37 or $(-5\delta_1)$ 33530·95 = ...55·50 are better. The latter makes mean limit .. 92·92 practically exact, and the separations 58·70, 249·00, 538, 680·65 all much improved.

$m = 6$. All the F are in good agreement with the calculated except for F_2. For this e.35617·10 = 28303·00, but too large. Also 28335·60, 28305·20 differ by 30·40 and $6\delta_1$ gives 29·58. Near F_1 32727·98, 32762·39 differ by 34·41 and $7\delta_1$ gives 34·37.

$-3\delta_1$ on the first or δ on the second give 32742·71 a better line for F as it makes the limit sum = 92·77 and gives better separations with $F_{3,4,5}$. There are clearly two sets with probable displacement in the f sequent. With the F_1 in the table would go better $(-6\delta_1)$ 33232·22 = ...62·39 for F_4 and $(-\delta_1)$ 33378·17 = ...83·21 for F_5. The linked F_6 agrees with (δ_1) 33427·04 = ...21·95.

$m = 7$. This presents several interesting points bearing on general theory. We may consider F_3 as correctly allocated since it differs only ·55 ($d\lambda$ = ·06) from the value calculated from the formula, but it is coincident with F_1 (7) of the 1864 series. Judging from the separations which are too small (except F_6) the observed F_1 is from 1 to 2 too large. This F = 28773·67 would seem to give some insight into the connection between sequent displacements and concomitant limit displacements or linkage attachments. Thus this line has relations with displaced limits with the two lines 28788·83 = $(-3\delta_1)$ F + ·43 and 28733·40 = (2δ) F − ·99 very close, but scarcely sufficiently so to exclude the probability of small sequent displacements. Further, it is linked forwards and backwards with all the three links e u.v as shown in the following scheme on the left. The 24638, 32912 form with F_1 an exact series inequality. Now a series inequality indicates that in the successive lines each is displaced from the preceding by the same amount, in this case about $15\delta_1$. The whole set may then be arranged as indicated in the right-hand scheme where X = 28771·80 = 28773·67 $(-k)$ and k denotes the displacement (? $15\delta_1$). The k may be the same for the different links within observation errors,* but probably not. If we take X as normal F_1 the other separations become normal also, but as they stand F_1 and **F** give the true mean limit. In other words they have the same sequent, or the sequent is modified in the same way for both. The normal $\mathbf{F}_1$ should be as much greater as the normal F_1 is smaller, and we do find such indications. The value given for $\mathbf{F}_1$ was found as the mean of two observed lines 32211·10 and 32213·80, supposed as ± displacements from the true line. If 32213·80 be taken as normal $\mathbf{F}_1$ it gives as mean limit with X ...92·80. The whole set of F and **F** in this order affords good illustrations of the remarks on *p* as to the general properties of these series. All the F lines show link connections. Thus F_2, $-(u-\cdot1)$, $u-\cdot1$; F_3, $+u-\cdot73$; F_4, $-(u-\cdot32)$, $u-2\cdot33$, $v-\cdot9$; F_5, $+v+1\cdot7$; F_6, $+u+\cdot2$.

---

* To make exact it would require the following:—7δ gives 1·84; 4δ, 1·05; 8δ, 2·10; 9δ, 2·36; 3δ, ·79.

		21456, 21462				$e.X$, $e.X(2k)$	
		$e-\cdot29\pm2\cdot38$					
24638				32912	$u.X$		$X(2k).u$
	$u+1\cdot87$		$u+1\cdot86$				
		F_1				$X(k)$	
	$v-\cdot70$		$v+2\cdot06$				
24346				33203	$v.X(k)$		$X(2k).v$
		$e-44\cdot0$					
		36083				$X(-k).e$	

In the preceding table the wave-numbers where deduced are entered as depending on one given line. But in nearly all cases these are substantiated also by links or other displaced limit lines. Although in an exhaustive treatment of the spectrum all these must be considered, this is not required for the illustration of general theory, and the notes do not go beyond $m = 7$. The number of the associated lines is remarkable. The proof of the existence of these lines depending on other *d* sequents as limits is so important, as a matter of general theory, that a large number have been adduced up to $m = 7$. To go beyond would be to overload the present communication with detail. The prevalence of the *u.v.* links over *e* may be noted in the F series.

In dealing with the 1864 series the mean of the common F and **F** limits for different orders was treated as giving the true value within a very small possible error of about $\xi = \cdot3$. It may be interesting to see how in the present cases the corresponding averages deviate from the values which have been established on the 1864 basis. The mean values found from the tables and expressed in terms of differences from the adopted ones are ·26, ·24, −·33, ·06, ·05, ·15 and justify the supposition of the small limit of error in ξ. The mean separations as found from the same 10 F lines deviate from the correct values by ·02, ·86, ·05, ·06, ·11. The large deviation ·86 is due chiefly to the uncertainty of the displacement in $F_3(3)$.

The preceding discussion is sufficient to show the excessive number and complication of the lines belonging to the F systems in this neighbourhood and that the complete problem has only been touched upon. As however the allocation of the extrapolated lines 16013 and 16044 as extreme satellites of D series based respectively on $D(\infty) = S(\infty)$ and $(-3\delta_1)S(\infty)$ is important, it will be well to consider shortly if corresponding F series can be indicated. The two lines referred to give the same *d* sequence, depending on a denominator $1+70\Delta_2$. This makes $d(1) = 35012\cdot30$.

Taking this as F (∞) and using the *f* (*m*) sequences as given by the formula (p. 412) it is possible to calculate the positions of the lines in question. The results are given down to *m* = 8.

m.	F.		35012·30.	**F**.	
	Calculated.	Observed.		Observed.	Calculated.
2	22652·26	...(3) 52·88	11·98	(...71·08)	47372·34
3	28092·94	(...93·54)	11·85	(...30·16)	41931·66
4	30594·43	...(1) 95·63	13·50	(...31·36)	39430·17
5	31948·38	(...47·00)	12·02	(...77·05)	38076·22
6	32763·63	...(1) 62·39	12·76	...(1*n*) 63·13	37260·97
7	32292·00	(...92·05)	12·14	(.. 32·23)	36732·00
8	33653·80	...(3) 53·04			36370·80

[*Note.*—δ_1 displacement on **F** (∞) gives a change 6·045.]

2. **F** requires sounding. (2) 35923·80.*e.v* = 47371·08.
3. $(-5\delta_1)$ (3) 28063·32 = ...93·54; **F** requires sounding. (δ_1) (1) 37508·25.*v* = 41930·21.
4. $(-2\delta_1)$ (2) 39419·27 = ...31·36.
5. $(-2\delta_1)$ (2*n*) 31934·91 = ...47·00; $(2\delta_1)$ (1) 38089·14 = ...77·05.
6. For **F**. (1*n*) 37263·13; also $(-6\delta_1)$ (<1) 37226·34 = ...62·88, or the same.
7. *v*.(1) 37720·05 = ...92·05; also $(-3\delta_1)$ (4) 33274·11 = ...92·24; (2) 32304·23.*v* = 36732·23.

The mean limit is 35012·37. The set form an additional test that the extrapolated lines 16013 really exist. It is curious to note that the even orders of F only show directly observed lines, whilst the odd show displaced limit lines. The **F** lines are far to the violet end and come into the observed region only when weakened by high order. The F lines are all linked to lines of higher frequency by the 1864 link, also for *m* = 4, 6 to lines of lower frequency. The same tendency is shown in **F** to lower frequency, any such to higher frequency lead to unobserved regions. This fact is important as showing that at least here the 1864 separation enters in the link relation, and not as a direct displacement on the limit.

To the F_3 (∞) limit corresponds a 1864 triplet series parallel to that originally considered. I have been able to follow it up in the same way as the foregoing as far as *m* = 26 at least. It accentuates the evidence for the displaced sets but as that is sufficiently supported by the results already discussed it would seem unnecessary to overburden the present communication with additional detail. Whenever the actual relations of the various displaced lines to one another are the subject of discussion these details will be of the first importance. A knowledge of these relations should be expected to throw a flood of light on the constitution of spectra, but this new question cannot be taken up here. It may be noted, however, that the first F and **F** lines of the triplets up to *m* = 10 are given in the table under F_3.

Ritz *Combinations.*—The results obtained enable us to test for additional series, associated with the name of **Ritz** who first pointed out their existence. They are represented by the expressions $n = p(1)-f(m)$, and $s(1)-f(m)$. Here $p(1) = S(\infty) = 51025·30$ and $f(m)$ is given by the formula on (p. 412). $s(1) = P(\infty) = 93178·69$ and should give a parallel series 42153·39 ahead, and therefore in the extreme ultra-violet. The results are for the sets p_1, p_2, p_3.

m.	Calc. $p_1-f(m)$.	$d\lambda$.	Observed.				
2.	38665·21	−·07	(1) 38666·56	**1784·06**	(2) 40450·62	**806·91**	(1) 41257·53
3.	44105·93	·18	(44102·34)	**1783·08**	(45885·42)	**815·66**	(46701·08)
4.	46607·42	·03	(46606·52)	**1784·28**	(48390·80)	**807·29**	(49198·09)
5.	47961·81	·06	(47960·42)	**1777·90**	+	**813·68**	(50552·00)

Notes.—$m = 2$. 1784·06+806·91 = 1777·90+813·07

$m = 3$. There is some ambiguity as to $F_3(3)$, which gives the calculated value. The deduced ultra-violet lines are from the observed triplet by the e link.

(7) 36788·24 **1783·08** (2) 38571·32 **805·66** (1) 39386·98.

Also

$e+v+(1)$ 32362·98 **1777·90** (<1) 34140·88 give 44105·08, 45882·98.

$u+(2)$ 39969·62 = 44102·8 with the ν_1, ν_2 lines in the ultra-violet, but an extra v sounder gives $u+v+(2)$ 37320·68 **813·93** (1) 38134·61 = 45881·86, 46695·79, the former being 1779·06 ahead of 44102·80.

$m = 4$. $e+(1)$ 39292·42 **1784·28** (1) 41076·70 = 46606·52, 48390·80 with $2e+(4)$ 34569·89 = 49198·09 which is 1777·90+813·67 above 40606.

$m = 5$. $2e+(2u)$ 33332·22 **1779·90+813·68** (2) 35923·80 = 47960·42, 50552·00; also $e+u+(1)$ 38292·32 **815·34** (1) 39107·56 = 49739·60, 50554·94; the former being 1777·79 above the calculated first line of 47961·81.

The Value of the Xenon Oun.—In the case of xenon the triplet disturbance appears to be small. The result has been obtained that $\Delta_1 = 24893 = 40\frac{3}{4}\delta$ and $\Delta_2 = 10999 = 18\delta$, giving respectively for δ the values 610·87, 611·05 or 610·94 from $\Delta_1+\Delta_2$. In the preceding 611 has been adopted for δ which is sufficiently close except where very large multiples of δ are in question. To obtain more accurate values of Δ_2 recourse must be had to the F and D sequences. Incidentally they have already been touched upon, but we are now in a position by a discussion of the whole material to attain a greater exactness.

Xenon, as has been seen, shows numbers of parallel displaced groups of D and F series, each of which gives data for the determination of Δ_2. Unfortunately those lines which might be expected to give the most accurate values are in the ultra-red beyond the observed region, and the process of extrapolation by the ν_1 separation, or by links, leaves a considerable margin of uncertainty owing to the 1780 modification of ν_1, and the well established, but not yet thoroughly understood, modifications of the e links, although in X this latter is not so marked as in Ag and Au. The material at disposal is—

From the D *Series.*

1. $[16013{\cdot}45] = 17791{\cdot}35-\nu_1 = 51025{\cdot}29+\xi-d(1+70\Delta_2)$ . . . (p. 404),
2. $[16044{\cdot}24] = 17829{\cdot}41-\nu_1-\nu_2 = (-3\delta_1)\ 51025{\cdot}29+\xi-d(1+70\Delta_2)$ (p. 409),
3. $19602{\cdot}66 = 51025{\cdot}29+\xi-d(1+79\Delta_2-\delta)$ (p. 403),
4. $[19623{\cdot}05] = 21400{\cdot}95-\nu_1 =$,, $-d(1+79\Delta_2)$. (p. 403),
5. $19942{\cdot}53 = (\delta)($,, $)-d(1+80\Delta_2-\delta_1)$ (p. 396),
6. $19989{\cdot}72 = ($,, $)-d(1+80\Delta_2)$. (p. 396),
7. $20021{\cdot}66 = (-3\delta_1)($,, $)-d(1+80\Delta_2)$. (p. 396).

From the F *Series.*

8. $[3010{\cdot}35] = 17638{\cdot}55-2e = 30725{\cdot}26+\xi'-f(1+90\Delta_2)$ . . . (p. 388),
9. $18380{\cdot}09 = F_3(2) = 30740{\cdot}17+\xi_1-f(2+89\Delta_2)$ . . . (p. 412),
10. $30725{\cdot}26+\xi' = d(1)$ sequent.

The relations between the ξ, ξ', ξ_1 have been already determined. $\xi = 1{\cdot}015\xi'+{\cdot}03+{\cdot}2q$ from the fact that (10) and the sequents of (6) are both d sequents (p. 407). Here q is the proportion of maximum error in (6). The ·03 may be supposed merged in this and $\xi = 1{\cdot}015\xi'$. Also $\xi_1 = \xi'+{\cdot}04$ from the fact that (10) and $30740{\cdot}17+\xi_1$ are both d sequences (p. 412). Within the accuracy attainable for our present purpose we may treat the ξ's as all equal. This was certainly not to be expected, for the limit $S(\infty)$, obtained by formula constants from S lines, is in general uncertain to a few units, whilst the ξ_1 and ξ', depending on limits found from $\frac{1}{2}$(F and **F**) should be very small.

The mantissæ conditions of the above give ($d\lambda = {\cdot}04p$)

*1. $769890+2{\cdot}5p_1-25{\cdot}27\xi = 70\Delta_2$ $\qquad \Delta_2 = 10998{\cdot}43-{\cdot}36\xi+{\cdot}036p_1$
*2. $769867+2{\cdot}5p_2-25{\cdot}27\xi = 70\Delta_2$ $\qquad = 10998{\cdot}10-{\cdot}36\xi+{\cdot}036p_2$
3. $868240+4{\cdot}6p_3-29{\cdot}73\xi = 79\Delta_2-\delta$ $\qquad = 10998{\cdot}16-{\cdot}37\xi+{\cdot}06p_3$
*4. $868846+5{\cdot}5p_4-29{\cdot}75\xi = 79\Delta_2$ $\qquad = 10998{\cdot}05-{\cdot}37\xi+{\cdot}07p_4$
5. $879711+4{\cdot}8p_5-30{\cdot}29\xi = 80\Delta_2-\delta_1$ $\qquad = 10998{\cdot}29-{\cdot}37\xi+{\cdot}06p_5$
6. $879853+4{\cdot}8p_6-30{\cdot}29\xi = 80\Delta_2$ $\qquad = 10998{\cdot}16-{\cdot}37\xi+{\cdot}06p_6$
7. $879855+4{\cdot}8p_7-30{\cdot}29\xi = 80\Delta_2$ $\qquad = 10998{\cdot}18-{\cdot}37\xi+{\cdot}06p_7$
*8. $989285+35dn-35{\cdot}4\xi = 90\Delta_2-\delta$ $\qquad = 10998{\cdot}84-{\cdot}39\xi+{\cdot}41dn$
9. $978842+16{\cdot}4p_9-120\xi = 89\Delta_2$ $\qquad = 10998{\cdot}22-1{\cdot}35\xi+{\cdot}18p_9$
10. $889326+\quad-30{\cdot}28\xi = 80\Delta_2+15\tfrac{1}{2}\delta$ $\qquad = 10998{\cdot}19-{\cdot}37\xi$

leaving out for the moment the extrapolated lines, indicated by *, and weighting No. 9 with three times the possible error of the others, the mean value of $\Delta_2 = 10998{\cdot}198 - {\cdot}37\xi$, the same as from No. 10 alone which is exact. They all, with the exception of No. 5, satisfy this within observation errors $< d\lambda = {\cdot}04$. No. 5 requires that the observation error shall be ·06A and the true wave-number 19942·47 in place of... 2·53. With this the ν_1 separation = 1775·69 and is brought into practically the exact $(\delta)\,\nu_1$ value (see p. 396) required, which is 1775·76. The outstanding ·07 ($d\lambda$ = ·017) would be attached to the strong second line of the doublet (10) 21717. This is in very striking support of the general argument. We have already seen good grounds for putting ξ a small fraction of the order ·25. To determine it with greater exactness a corresponding mantissa differing from the above by considerable multiples is necessary: *e.g.* with mantissa of order ·5 the coefficient of ξ is $15{\cdot}4\xi$. The differences equated to Δ_2 multiples would then give an equation to find ξ in which the error term would have little effect. We get this different ξ coefficient in No. 9, but it is due to an order 2 in which the effect of an error is multiplied to the same extent. The extrapolated lines do not help us as their limits of error are too large. On the contrary the argument enables us to determine their values more correctly: *e.g.* in No. (2) the error is dependent as the line 21400 from which the line is extrapolated. To make the multiple correct requires $p = 2{\cdot}7$, $dn = {\cdot}42$. This reduces the observed $\nu_2 = 809{\cdot}53$ to $809{\cdot}11$. It is supposed modified by a δ_1 shift on the sequent which here produces a change of 6·3 pointing to an original $\nu_2 = 809{\cdot}11 + 6{\cdot}03 = 815{\cdot}14$, practically exact. Applying the method to Nos. 3, 9 gives

$$
{\cdot}98\xi = {\cdot}06 + {\cdot}18p_3 - {\cdot}06p_3
$$

$$
\xi = {\cdot}06 + {\cdot}18p_9 - {\cdot}06p_3 = 06 \pm {\cdot}24
$$

with

$$
\Delta_2 = 10998{\cdot}14 - {\cdot}064p_9 + {\cdot}043p_3 = 10998{\cdot}14 \pm {\cdot}10
$$

But the preferable choice is to use the fact that (10) is the limit to (9), the same value of ξ must enter, and the result depends only on the observation error.

The result is now

$$
{\cdot}98\xi = {\cdot}03 + {\cdot}18p_9, \qquad \xi = {\cdot}03 + {\cdot}18p_9, \qquad \Delta_2 = 10998{\cdot}187 - {\cdot}06p_9.
$$

Thus with maximum error $d\lambda = {\cdot}04$ maximum uncertainty in Δ_2 is ±·06, but the line (9) is a good one for measures and the probable error will not exceed ·02. Hence as the definitive value $\Delta_2 = 10998{\cdot}18$ is probably within ·03 and certainly within ·06. Hence

$$
\Delta_2 = 10998{\cdot}187 \pm {\cdot}03, \qquad \delta = 611{\cdot}0104 \pm {\cdot}0017.
$$

The value of δ obtained from the ν_1 displacement $= 610{\cdot}87 + {\cdot}76d\nu_1 - {\cdot}04\xi$, to make these the same requires $d\nu_1 = {\cdot}19$, $\nu_1 = 1778{\cdot}09$. This is possible though not probable. We cannot say definitely here therefore as in Kr that the triplet modification produces a slight difference in the deduced δ. Here the difference ·14 would correspond to a difference in mass of 27 electrons.

Radium Emanation.—The emanation does not apparently produce the two spectra exhibited by Kr and X. The measurements in the spectrum are very scanty compared with those in the latter. We have early rough determinations by Ramsay and associates.* More accurate and complete by Rutherford and Royd†, and later by Watson.‡ In order to diminish the absorption by the electrodes Ramsay also used copper instead of Pt electrodes and found a number of new lines, the majority of which have not been seen by succeeding observers. They have generally been explained as due to contamination by xenon as they lie close to X lines within their errors of observation. At first sight this explanation would seem to be very natural, but Ramsay was confident that there was no such contamination. I am inclined to suspect that the opinion that these lines belong to X is too hasty. As is well known Baly found quite a large number of lines in Kr and X coincident within his observation errors, which indeed were much smaller than those in any measures yet made in RaEm. Now as a fact those suspected lines of Ramsay and Cameron's are also very close to these Kr lines. A strong argument also is this. There are a number of strong lines undoubtedly belonging to the RaEm spectrum, and observed by both Rutherford and Royd and by Watson, which are also near strong X lines, yet separated so far from them, that if Ramsay and Cameron had had X in their tube they must have seen them and RaEm lines as double, one due to X and the other to RaEm. Compare for instance the following lines:—

RaEm.			X.	Kr.
C. and R.	R. and R.	W.		
(5) 4681	(10) 4680·92	(9) 4681·01	(5) 4683·76	(4) 4680·57
(10) 4626·5	(8) 4625·58	(10) 4625·66	(15) 4624·46	
(10) 4605	(4) 4604·46	(8) 4604·58	(10) 4603·21	
(3) 4578·5	(7) 4577·77	(8) 4578·0	(6) 4577·36	(6) 4577·40
(8) 4463·5	(7) 4459·3	(10) 4460·0	(20) 4462·38	(1) 4463·88
(3) 4189	(4) 4187·97	(5) 4188·2	(10) 4193·25	
(6) 4114	(6) 4114·62	(6) 4114·71	(7) 4116·25	(1) 4113·90

I therefore included these Ramsay and Cameron lines in the purview, with the result that a considerable number were found to fall in with places in which they are required by known spectrum laws. It must be remembered that the appearance of the Em lines varies very much in relative intensity with different observers (*cf.*, for instance 4604, 4460 above) that some appear early and then disappear, that others come in after the emanation has stood for a few days, and further that the copper electrodes, which extended the useful duration of the tubes, would probably have some effect on the nature of the emitting sources in the gas. To account for this, the suggestion might be thrown out that the activity of the emanation would by itself ionize the molecules of the gas, and that especially the α-rays would ionize in a different and more drastic way than the ordinary cathode or vacuum tube ionization. That with time the γ-rays from the active deposit might ionize in again a different way and produce again new lines. One would expect that the self-effect—as it may be called—is so drastic that it destroys those configurations which should give the red spectrum analogous to that in the other gases. It is a fact, as I hope to show, that the spectrum, so much as there is of it, is decidedly of the jar, or blue kind.

---

* Ramsay and Soddy, 'Roy. Soc. Proc.,' vol. 73, p. 346 (1904); Ramsay and Collie, *ibid.*, vol. 73, p. 470; Cameron and Ramsay, *ibid.*, A, vol. 81, p. 210 (1908).

† Rutherford and Royds, 'Phil. Mag.' (6), vol. 16, p. 313 (1908): Royds, *ibid.*, vol. 17, p. 202 (1909).

‡ H. E. Watson, 'Roy. Soc. Proc.,' A, vol. 83, p. 50 (1909).

The degree of accuracy of the observations is not of the best. ROYD claims an accuracy of 0·1A. The spectrum was obtained by a concave grating of 1 metre radius and extended from 5084 to 3005, with some additional lines by a prism spectrograph, subject to errors of ·5A. WATSON's lines extended from 7057 to 3867 with several new lines. His degree of accuracy is probably about the same as that of ROYD. In the following we shall treat the maximum errors as ·2A except where lines are only given to the nearest unit.

The extent of the spectrum observed is too restricted to expect to find more than the S (2) and D (1) lines, and even in the case of S (2) the $S_2(2)$ and $S_3(2)$ may be in the violet where only glass apparatus was used. Further, there is the added disadvantage that the links are so large that they can stretch from the unobserved ultra-red to the unobserved ultra-violet, and consequently can only act as sounders for lines so far in the ultra-violet that an *e* link lands within the visible red. The F lines should be expected to lie wholly in the observed region, and this must be the chief guide in the unravelling of the series relations.

As a preliminary and definite starting point, we have the value of the oun as calculated from the atomic weight. But here also there is some uncertainty. The value of HÖNIGSCHMIDTS' determination of the atomic weight of Ra, 225·97 is now generally accepted as close to the real value, as against the earlier value of 226·4. This makes the atomic weight of the emanation to be 222 to 222·4. These two give values of $\delta = 361{\cdot}80w^2$ as between 1783·1 and 1789·5, with the probability that it is close to 1783. The uncertainty in the value of the constant 361·80 will not affect this.

An examination of the spectrum for constant separations shows a large number of triplets with ν_1 in the region 5371 to 5383 and ν_2 at 2671 and less. Further, the higher values appear in sets which show inverted order of intensities. This suggests that the lines belong to D satellite systems and that the separations about 5383, 2671 belong to the modified ν_1, ν_2 which D satellites have already shown in Kr and X, whilst 5371, 2641 or thereabouts belong to the normal separations depending on displacements in the S(∞) alone. Moreover, another very frequent separation is 5631, connected with other sets as triplets with a ν_2 in the neighbourhood of 2800. This at once suggests the analogue of the 1864 F series of X.

A first quite definite starting point, from the material at disposal, is found by a search for lines of the F and **F** type, or the twin A±B sets. The limit A belongs to a D sequence, which from analogy with Kr and X should be expected to be of the order $n = 30000$. Now in the observed spectrum there is a long gap between 27671 and 32031, within which such limit must be. That no lines should be found near this limit is to be expected. If, however, such double sets exist we should expect to find sets of lines with exactly the same separations on either side of this gap. Unfortunately there are only four lines on the violet side, but one such set is found. They are

(1) 26669·23 **897·13** (2) 27566·36

(0) 33259·05 **896·91** (0) 32362·14

The corresponding limit should be the mean of either of the two corresponding lines, viz., 29964·14 or ...4·25, say 29964·20±·05. The possible observed errors in these lines are not large and the practically exact equality of the two separations is strong evidence of the reality of the suspected connection. But any doubt on this point must be removed when it is noted that by RYDBERG's tables, the separation 897 is that due to two denominators 5·78, 6·78, whilst if the denominators are calculated using 29964 as limit the same values are found. The two results are quite independent. By a further use of RYDBERG's tables it is possible to find approximate positions for other lines of the F system, the **F** being quite beyond the observed region in the ultra-violet. Such lines are found at (5) 15846, (4) 22310·4, (8) 25171 for $m = 2, 3, 4$. Again, connected with 15846, are (3) 21488, (6) 24296·9 giving separations 5642, 2809. The lines 22310, 21487 are due to C. R.'s copper electrodes and are subject to considerable possible errors $d\lambda = 1\text{A}$, $dn = 5$, or even more. These strikingly correspond to the 1864 sets in XF.

The limit 29964·20 must be very accurate and subject only to any systematic errors in R. and R.'s measurements. This is shown by the exactness of the observed separations in sets so far removed from one another as 26669 and 32362. Using this limit with the lines 15846, 26669·23 for $m = 2$ and 5 the calculated formula is

$$n = 29964{\cdot}20 - \mathrm{N} \Big/ \left\{ m + {\cdot}757457 + \frac{{\cdot}059446}{m} \right\}^2 .$$

The two lines used may be regarded as having possible errors $dn = 2$ and ·7 respectively and any consequent errors in the constants will scarcely affect the calculated frequencies for $m > 2$. These for $m = 3$, 4, 6 are 22277·28, 25148·60, 27569·40. The last gives $d\lambda$ = ·40. No lines correspond to the others, but 22211 observed by C. and R. with Cu electrodes and (15) 25107·14 (R.), (9)...6·64 (W.) are respectively 66·28 and 41·56, (41·96) less. Now a δ_1 displacement on the limit gives 13·96, so that these correspond within error limits to a $(5\delta_1)$ F (∞), $d\lambda = -\cdot5$, and $(3\delta_1)$ F (∞), $d\lambda = -\cdot06$, and stand in some analogy to what has been seen to happen in X. The intensity of the second line however would seem very great for $m = 4$. In support of the C.R. line with its large possible error is (2) 22196·92 (Royd) at 13·08 behind it or an extra δ_1. It would make 22196 = $(6\delta_1)$ F, and consequently F = 22280·68, *i.e.*, $d\lambda = -\cdot7$. If the 5640 lines exist, they all lie outside the region of observation except in the case of the first line 15846 where the corresponding triplet set is found (see above). However, anticipating the value of the links obtained later and using them as sounders, viz., e = 23678·3, u = 11191·8, v = 13680·6 with some uncertainties we can test for their existence. Taking the observed lines 22196·92, 22211 we expect lines about 27838, 27858, only to be observed if strong. There are none, but there are (3) 14166 (W.) **2816** 16982·0 (C.R.) which with the v link give 27846·6 ± 1·9 + dv, 30662·6 ± 3 + dv, separation = 2816 ± 5. These suggest the triplet set (δ_1) F, viz.,

22196·9	**5649·7**			
22210·9	**5635 + 1·9p + du**	27846·6 + 1·9*p* + *dv*	**2816 + 3p′ − 1·9p**	30662·6 + 3*p*′ + *dv*

The results of sounding are indicated in the following, where, since e, u, v are only approximately known, their values are supposed corrected by de, du, dv. The letters after wave-lengths refer to observers:—

$m = 3$. $(5\delta_1)$ F.

Outside		(3) 14166 (W.).*v*		16982 (C.R.).*v*
22196·9	**5649·7**			
22210·9	**5635 + 1·9p + dv**	27846·6 + 1·9*p* + *dv*	**2816 + 3p′ − 1·9p**	30662·6 + 3*p* + *dv*

With the calculated value of F_1 the first separation is 3·5 larger. The first line is δ_1 displacement on 22210. If the former does not belong to the system, we may take C.R.'s direct line which has possible error of 4 or 5.

$m = 4$. $(3\delta_1)$ F.

1.	Outside		$(2\delta_1)$ 19583 (C.R.).*u*		$(2\delta_1)$ 22398 (C.R.).*u*
	25107·14	**5640·6 + 4p + du**	(30747·8) + 4*p* + *du*	**2814·2 + 5p′ − 4p**	(33562) + 5*p* + *du*
2.	*v*.(0) 15136 (W.).*e*		*v*.(0) 20784·31 (W.).*e*		*v*.(0) 23598·2 (W.).*e*
			v.$(-\delta_1)$ (1) 20799 (C.R.).*e*		
	(25134·3)	**5648·31 + de − dv**	(30782·6)	**2814·9**	(33596·5)
3.	[25148·60]	**5648**	*v*.(1) 20799.*e*	**2816·0**	*u*.(1) 21126·0.*e*

No. (2) is close to $(-2\delta_1)$ 25107 or (δ_1) F_1. It should be noted that the C.R. copper line is a $(-\delta_1)$ displacement on that observed by WATSON, and that the latter was only observed after the emanation had been standing two days. No. 3 gives the calculated values.

$m = 5.$

1.	Outside		(1) 21126·0 (W.).*u*		23948 (C.R.).*u*
	(1) 26669·23	**5648·6+du**	32317·8	**2822+6p**	35139+6*p*
2.			(2) 18641·4 (W.).*v*		(6) 21458·86 (W.).*v*, (6·56) (R.R.)
	26669·23	**5652·8+dv**	32322·0	**2817·4**	35139·46
3.			*u*.(1) 19842·4 (W.).*e*		*u*.(4) 22636 (C.R.).*e*
	26669·23	**5659·7+de−du**	32328·9	**2793·6+6p**	35122·5

These would seem to show in Nos. 2, 3 sequence displacements; 2δ would here produce a change of 4. In (2) it is displaced in the second and continued in the third. In (3) it is displaced only in the second. 21126 acts as a sounded line for $F_3(4)$ and $F_2(5)$. Both cannot be real or the line is double.

$m = 6.$

1.			(1) 22015 (C.R.).*u*
	(2) 27566·36	**5640·4+6p+du**	33206·8+6*p*
2.			(3) 19527·8 (W.).*v*
	27566·36	**5642+dv**	33208·4

There is a C.R. copper line (2) 16386±3 which with the *u* link gives 27577±3 and may be $(-\delta_1)$ F (6).

The two last observed lines have been included in order to show that sounding is justified by them. The calculated line for $m = 7$ is 28155·7, with error probably within 3 or 4 units. There is 16982 (C.R.) which with *u* sounder gives 28173·8±3 and may well correspond to (δ_1) $F_1(7)$. Also 22636 (C.R.).*u* gives 33827·8 which is 5654 ahead and may be (δ_1) $F_2(7)$. But this line with an $e-u$ link also gives $F_3(6)$. It should be noticed how the lines observed by CAMERON and RAMSAY with copper electrodes come in to fill parallel and displaced lines where they seem called for.

The corresponding **F** lines for the orders above the observed ones are in the ultra-violet, but evidence for their existence is given by sounding. In what follows, the value of $\mathbf{F}_1$ calculated from F_1 is enclosed in square brackets.

$m = 2.$ [44082·4±2.]

1.	(δ_1) (6) 21712·09 (R.).2*u*		$(3\delta_1)$ (0) 27389·61 (R.).2*u*	
	44081·7+2*du*	**5649·6**	49731·3+2*du*	
2.	(3) 16725·1 (W.).2*v*			(8) 25170·10 (W.).2*v* ?
	44086·3+2*dv*	**8445 = 5645+2814−13·96**		52531·3+2*dv*

In No. 1 the sounded line $\mathbf{F}_2$ is on the verge of the observed region. In No. 2 the sounded is about 4 too large, and there appears a δ_1 displacement on the last. The u, v links themselves are too short to reach.

$m = 3$. [37650·6 from Calculated F, 37717 from Observed $(5\delta_1)$ F.]

(3) 23973·69.v		(0) 15944 (W.).$2v$		(2) 18739 (C.R.).$2v+4p$
37654·29$+dv$	**5651+dv**	43305·2$+2dv$	**2795+4p**	46100·2$+2dv$

Here the reproduced line refers to the normal line calculated from the formula. The v link is too short for the second and third lines, $2v$ reaches it, but $2v$ on the first would require a reference line in the ultra-red. Also 15944 is possibly $(-7\delta_1)$ F_1 (2).

$m = 4$. [34779·80.]

(3) 23584 (C.R.).u (δ_1) (0) 23598·2 (W.).u		$(-2\delta_1)$ (3) 16725·1 (W.).e
34776·0	**5655·2+de−du**	40431·2

Again note a C.R. copper line supported by a W. and R.R. displaced δ_1 line

$m = 5$. [33259·17.]

		(2) 25213·7 (W.).v		(0) 18024 (W.).e
33259·05	**5635·25+dv**	38894·3	**2808·0+de−dv**	41702·3

Here appears the frequent 5635. It is 5649−14, that is, there is a limit displacement of δ_1 in the second line and an extra one in the third, making the second separation 2816.

$m = 6$.

The same links cannot serve as sounders for all three lines.

For $\mathbf{F}_1$ is (2) 211725·4 (W.).u = 32364·34$+du$ as against calculated 32362·14.
For $\mathbf{F}_2$ is v.(2) 16816·5 (W.).$e.u$ = 38004·8$+de+du-dv$.
For $\mathbf{F}_3$ (4) 18448 (C.R.).$2u$ = 40831·6$+2du$.

These give separations 5640·5, 2826·8, where the sum has the normal value. In $(-\delta_1)$ (3) 23973·69 (W.).u = 35178·45 which is 2816·2 ahead of $\mathbf{F}_1$ we have the completion of a mesh with the other three lines.

The foregoing discussion has shown: (1) that this F set belongs to the 1864 type discovered in X; (2) that the usual displacements are present and that the calculated value per oun—13·96—satisfies all the numerical relations formed; and (3) that the lines observed with copper electrodes by CAMERON and RAMSAY seem to belong specially to parallel series to this set. We could feel complete confidence in the allocation of the lines were it not that the α constant in the formula is positive, and that the line $m = 6$ is not reproduced more closely. The absence of direct representatives for $m = 3$, 4 is not surprising as their limit displaced values are certainly observed and the change is in full agreement with what takes place in the other elements.

We shall assume in what follows that the preceding allocation is correct, in other words the limit is $29964{\cdot}20 + \xi$, the line for $m = 2$ is $15846 + 2{\cdot}5p$, and that the series belongs to the F type. In that case the mantissa of the limit and of the sequent are both multiples of the oun. These mantissæ are respectively $913165 - 31{\cdot}92\xi$ and $787174 + 246p - 98{\cdot}7\xi$. As both are oun multiples, so must be their difference. This difference is

$$\begin{aligned}125991 - 246p + 68\xi &= 70\tfrac{1}{2}(1787{\cdot}10 - 3{\cdot}49p + {\cdot}97\xi)\\ &= 70\tfrac{3}{4}(1780{\cdot}75 - 3{\cdot}49p + {\cdot}97\xi)\end{aligned} \qquad (1)$$

In which if WATSON is correct to nearest unit p is equally probable between $\pm{\cdot}5$.

It is very unfortunate that here we have to deal with two uncertainties not generally met with, viz., on the one hand the uncertainty as to the real value of the atomic weight, and on the other the magnitude of the possible observation error in the fundamental wave-length, which WATSON has only measured to the nearest Ångström. If this had been ·1, *i.e.*, $p = {\cdot}1$, the above result would show that since δ lies between 1789 and 1783, the multiple must be $70\tfrac{1}{2}$ without any doubt, and consequently δ in the neighbourhood of 1787. The value of ξ is so small, that its term will not affect our present reasoning. We have, however, to allow for this uncertainty and a value of $p = -{\cdot}7$ makes the second multiple $= 70\tfrac{3}{4} \times 1783{\cdot}2$ with a possible δ. In this case the first multiple gives $70\tfrac{1}{2} \times 1789{\cdot}55$, or δ just on the improbable limit and it might be excluded. The result, therefore, is

Equally possible, $p < {\cdot}5 > -{\cdot}5$, multiple $= 70\tfrac{1}{2}$ and δ between 1785·3 and 1788.
Improbable, but perhaps possible, $p = -{\cdot}7$, multiple may be $70\tfrac{3}{4}$ and $\delta = 1783{\cdot}1$.
Very improbable, $p = 1$, multiple $70\tfrac{1}{2}$ and $\delta = 1783{\cdot}6$.

But also the limit and sequent mantissæ must also be oun multiples, now

$$\begin{aligned}913165 - 31{\cdot}92\xi &= 512\,(1783{\cdot}52 - {\cdot}062\xi)\\ &= \ldots\\ &= 511\,(1787{\cdot}016 - {\cdot}062\xi)\end{aligned} \qquad (2)$$

$$\begin{aligned}787174 + 246p - 98{\cdot}7\xi &= 441\tfrac{1}{2}(1782{\cdot}95 + {\cdot}55p - {\cdot}22\xi)\\ &= \ldots\\ &= 440\tfrac{1}{2}(1787{\cdot}00 + {\cdot}55p - {\cdot}22\xi)\end{aligned} \qquad (3)$$

It might occur to the reader that the last should be a multiple of Δ_2. But if the series is the analogue of the 1864XF, to which the foregoing argument has pointed, it should have a line of order $m = 1$ (n about $= 3260$). This should show $M(\Delta_2)$.

The limit condition is independent of p and can only be modified by ξ in the second decimal place. It gives quite definitely six possible values for δ, viz., 1783·52, 1784·39, 1785·26, 1786·13, 1787·01, 1787·87 with multiples 512 diminishing by $\frac{1}{4}$ to $510\frac{3}{4}$. Of these the following can be satisfied by the mantissa difference condition (1)

1783·5 by $p = -·8$, multiple $70\frac{3}{4}$, not probable.
1783·5 by $p = 1$, multiple $70\frac{1}{2}$, very improbable.
The last four by $p > -·25 < ·5$, multiple $70\frac{1}{2}$, equally probable.
The others by $p > ·5 < 1$, multiple $70\frac{1}{2}$, improbable.

If Watson's readings are really to the nearest unit, $p = \pm·5$. This probable consideration would largely reduce the limits of uncertainty. It would in conditions (1) exclude the second and with multiple $70\frac{1}{2}$ give 1787·0 with $p = 0$, 1786·1 with $p = ·3$, 1785·3 with $p = ·5$. Conditions (3) would then of these give 1787 with $p = 0$, multiple $440\frac{1}{2}$, 1786·16 with $440\frac{3}{4}$, 1785·27 with 441. All of these have equal probability, but they exclude the 1783 based on Hönigschmidts' atomic weight. The lowest value 1785·3 would make the atomic weight = 222·15 ± ·02, and that of Ra = 226·15 as compared with Hönigschmidts' 225·97.

Before passing from this series it will be important to get as close an estimate as possible of the two separations. Regarded as oun displacements on the limits they should give some further data for the determination of the oun—or, vice versa. The separations given in the sounding operations above are here collected. Errors from Watson's, or Baly's observations will not amount to more than a few decimals at the outside.

F.			F.		
1.	$\left.\begin{matrix}5649\cdot7\\35\cdot7\end{matrix}\right\}+1\cdot9p+dv$	$2816+3p'-1\cdot9p$	9. 10.	$5649\cdot6$ $50\cdot9+dv$	 $2795+4p$
2.	$40\cdot6+4p+du$	$14\cdot2+5p'-4p$	11.	$55\cdot2+de-du$	
3.	$48\cdot31+de-dv$	$14\cdot9$	12.	$35\cdot25+dv$	$2802\cdot1+de-dv$
		$16\cdot0-du+dv$	13.	$36\cdot6$	$33\cdot0$
4.	$48\cdot6+du$	$22+6p$			
5.	$52\cdot8+dv$	$17\cdot4$			
6.	$59\cdot7+de-du$	$793\cdot6+6p$			
7.	$40\cdot4+6p+dv$				
8.	$42\cdot0+dv$				

The ν_1 cluster round 5649, 40, 35 and the ν_2 around 2816.

In ν_1 the 49 and 35 differ by 14 and are clearly due to a δ_1 displacement in the limit. We will consider the exceptions in order. In (2) there is an uncertainty $4p$. If $p = -1$ we get, with other errors close to 35, in this case $\nu_2 = 18·2+5p'$ and a small error in p' brings it to the 16 neighbourhood. In (4) the uncertainty $6p$ reduces again ν_2 to the 16. In (5) ordinary errors bring ν_1 to the 49 value. In (6) modification of 10·7 in the middle line brings ν_1 to 49 and $p = ·25$ brings ν_2 to 16 or $\nu_1+\nu_2$ = normal. In (7) the error allows 35. No. 8 does not seem amenable and may not therefore be a real connection. In (10) ν_1 is a normal 49 with error 2 and ν_2 with $p = 1$ is $00 = 14-14$ or the oun displacement in the last line. In (12) also the oun displacement in the last line makes $\nu_2 = 16$. No. 13 can be explained by the oun displacement on the middle line alone. They can all then with the exception of one be explained by $\nu_1 = 5635$ or 5649, and $\nu_2 = 2816$ within limits of about ± 1. The consideration of their source is postponed until the D sequents have been touched upon (p. 441).

An examination of the spectral list shows values of ν_1 varying between 5361 and 5388 with ν_2 from 2639 to 2685. In many cases two lines are separated by about 5371 and a line approximately midway with separations $2680\pm$. Now in the other gases the appearance of such $\frac{1}{2}\nu_1$ separations is so common as to be almost the rule. This lends weight on the one side to the assumption that $5371\pm$ is ν_1 and on the other that 2680 is not ν_2 but corresponds to $\frac{1}{2}\nu_1$. Values therefore in the neighbourhood of 2686 may be left out of account in the search for ν_2 which is always less than $\frac{1}{2}\nu_1$.

There are a considerable number of separations in the neighbourhood of $5380\pm$ which again suggest the analogous modified ν_1 related to the D satellites in Kr and X. In those cases the explanation was adopted that at least in the main they arose from displacements of small multiples of the oun in the sequence term. No exact value of these displacements can be obtained until the value of the limit itself is known, but it is possible to arrive at an approximate estimate by employment of a value of the limit which may be a few hundreds wrong. Such an estimate will be of great value in guiding our search.

The values of $S(\infty)$ for A, Kr, X are respectively 51731, 51651, 51025. From analogy, that for Ra, Em would be in the neighbourhood of 50500, say $50500+\xi$. With ν_1 in the neighbourhood of 5371 ± 1 calculation shows that a displacement of δ_1 will produce a change in the limit of $30{\cdot}86+{\cdot}0011\xi$ and in the ν_1 of 4·72. The D sequence terms vary with each satellite set. One such sequence set has already been found in the F series already discussed, viz., 29964. Here the oun displacements produces a change in n of 13·96. But higher values than this for $d(1)$ sequents are to be allowed for up to even 31000. For 30500 the oun displacement produces a change of 14·88. It is safe therefore to take the oun displacement on the sequent as producing a change varying from 14 to 15.

If now the $5370\pm$ separations be analysed it will be found that three lie at $71{\cdot}4\pm{\cdot}4$, eight at $74{\cdot}2\pm{\cdot}6$, five at $80\pm{\cdot}5$, six at $84{\cdot}2\pm 1$ and a group at 88 with none between. There is a clear majority about 74 with a set at 88, or about 14 ahead (*i.e.*, $-\delta_1$ on the sequent) and a few about 61, or 14 behind. Whilst the 71·7 and 78 sets suggest the δ_1 displacements in the limits, definitely seen to exist in Kr and X.

For the ν_2, the oun changes on the sequent are the same and on the limit are about half the previous, say, about 2 or 3. In this case are found a majority near 2649 with some about 14 on either side, and another few about 2652.

The material at disposal then goes to show that the S separations are near 5371, 2649 for some definite limit and 5374, 2652 for another whose limit mantissa is one oun less.

Arguing from analogy with the successive spectra of A, Kr, X we should expect to find in the observed region only lines corresponding to D (1) and S (2). The D (2) lines would be considerably shorter than the last observed line $n = 33259$. The S (1) would be in reversed order with $S_1(1)$ near -42100 or $S_3(1)$ near -34000, the first absorbed by the glass apparatus used. We should expect D (1) lines up to the longest observed ($n = 14166$) with $D_{11}(1)$ lines down to at least $n = 21000$, and showing the same kind of modified separations as in previous cases, and taken account of above. $S_1(2)$ should be about 25600 with $S_2(2)$ about 31270, just on the observed boundary and $S_3(2)$ quite beyond about 33900. This absence of S separations is the reason why it was so difficult above to obtain accurate values of ν_1, ν_2 for a definite limit. By themselves therefore the material is hopelessly inadequate to determine the $S(\infty)$ limit, the values of Δ_1, Δ_2, or of the various links. We have only five possible—or three probable—choices for δ to 5 significant figures, and also the value of one $d(1)$ sequent correct to a few decimals, with estimates of the F and S triplet separations. The only method of attack then seems to be an indirect one, to tabulate the sets of lines giving the triplet separations, to try to distinguish between those related to D and S systems, to obtain as close a value as possible of the ν_1, ν_2, to determine some of the satellite separations, and from these last to attempt to find the corresponding F series with the same constant separations. These F series ought in each set to consist of several orders (m) at least, as the $F(\infty)$ all lie in the observed although badly observed region. The observed separations and the values of ν_1, ν_2 combined with the approximate value of the oun may enable a determination of the important constants $S(\infty)$, Δ_1, Δ_2 to be arrived at.

We shall take then $5371 + d\nu_1$ $2649 + d\nu_2$ and a set about 3 larger for the values of ν_1, ν_2, the two sets belonging to two limits, relatively displaced by an oun, and both giving the same values of Δ_1, Δ_2. In the above the $d\nu_1$, $d\nu_2$ will probably not be greater numerically than 1. The calculations will be made with the 5371 set and the conditions applied $d\nu_1 = \pm 1$, &c., and $d\nu_1 = x \pm 1$, $d\nu_2 = y \pm 1$ where x, y are the changes produced by the oun displacement. We start the first approximation by taking $S(\infty) = 50500 + \xi$ where ξ may amount to several hundreds and in which x, y are of the order 4·7, 2·2. The denominators of the S_1, S_2, S_3 limits with their differences calculated from these are

$$\begin{array}{ll}
1{\cdot}473697 - 14{\cdot}591\xi & \\
 & 72624 - 2{\cdot}053\xi + 12{\cdot}54\, d\nu_1 \\
1{\cdot}401073 - 12{\cdot}538\,(\xi + d\nu_1) & \\
 & 32076 - {\cdot}839\,(\xi + d\nu_1) + 11{\cdot}70\, d\nu_2 \\
1{\cdot}368995 - 11{\cdot}699\,(\xi + d\nu_1 + d\nu_2) &
\end{array}$$

These differences may be written

$$\begin{array}{llll} 72624 = 40\tfrac{3}{4}\,(1782{\cdot}18 - 050\xi + {\cdot}31 d\nu_1) & & 32076 = 18\;\;(1782{\cdot}00 - {\cdot}047\xi - {\cdot}047 d\nu_1 + {\cdot}65 d\nu_2) \\ \quad = 40\tfrac{1}{2}\,(1793{\cdot}18 \ldots \quad ,, \quad\;) & & \quad = 17\tfrac{3}{4}\,(1807{\cdot}1 \ldots \qquad ,, \qquad\;) \end{array}$$

The first consequence to be drawn from these numbers is that as δ is to be of the same order of magnitude as calculated from Δ_1 and Δ_2 it can only take place in the first arrangement since the co-efficients of ξ (the only variable capable of large values) are practically equal. In other words $\Delta_1 = 40\frac{3}{4}\delta$, $\Delta_2 = 18\delta$, the same multiples as in X. This is one quite definite conclusion of importance.

But further we know that δ must be one of the following sets 1783·5, 84·4, 85·3, 86·1, 87·0, of which the first two though perhaps possible are improbable and the last three equally probable. We will state the results for the first, third, and fifth, the others being intermediate. The values of $d\nu_1$, $d\nu_2$ being uncertain to about unity, the values of ξ can only be estimated to the same order. Taking Δ_1, in order to produce the required values of δ, ξ must have the respective values of -27, -64, -100 within a few units, and the numbers inside the brackets for Δ_1, Δ_2, become

$$\begin{array}{lll} 1783{\cdot}53 - {\cdot}050\xi + {\cdot}31 d\nu_1 & \qquad & 1783{\cdot}27 - 047\xi - {\cdot}047 d\nu_1 + {\cdot}65 d\nu_2 \\ 1785{\cdot}38 - \ldots \quad ,, & & 1785{\cdot}00 \ldots \qquad ,, \\ 1787{\cdot}18 - \ldots \quad ,, & & 1786{\cdot}7 \ldots \qquad ,, \end{array}$$

where now ξ cannot exceed a few units. The respective limits become 50473, 50436, 50400. One of these is close to a S limit. There is another corresponding to the $-\delta_1$ displacement about 31 ahead and giving $\nu_1 = 5375$ about. We already have an accurate value of one d sequence, viz., the F $(\infty) = 29964{\cdot}20$. Combined with the limit this should give a D line $= \mathrm{S}(\infty) - 29964{\cdot}20$, *i.e.*, near 20509, 20472, or 20435, or the 31 higher values. We only find the following possible lines, viz. (2) 20515 and (4) 20438·8 (W.), (0) 20446·38 (R.R.), 20473, the first and fourth being C.R.'s copper electrode lines. The first is one of a chain of 2680 separations, or $\frac{1}{2}\nu_1$ stretching back to 15136 and by 2689 to the strong line (10) 23204·6 (R.) or 7·57 (W). It has none of the signs of a D line. The line 20438 is one of the strongest in this neighbourhood has no ν_1 separation link, but 5625·8 back to 14813±2 (W) to which also 20446 has the (δ_1) 5649 link, viz.,5633. Again the 14813 shows a D modified ν_1, ν_2 set, viz.,

$$\begin{array}{llllll} (0)\ 14813\ (\mathrm{W}) & \mathbf{5382{\cdot}71} & (4)\ 20195{\cdot}71\ (\mathrm{W}) & \mathbf{2672{\cdot}41} & (5)\ 22868{\cdot}12\ (\mathrm{W}), \\ & \mathbf{5385{\cdot}91} & (0)\ 20198{\cdot}91\ (\mathrm{R}) & \mathbf{2669{\cdot}16} & (1)\qquad {\cdot}07\ (\mathrm{R}). \end{array}$$

Hence here we have all the signs of a D_{11} line in 20438 with the above as a satellite set, at the F separation (56...) behind. Now it is a very striking fact that the measures of W. and of R.R. for these lines not only differ by amounts larger than can be attributed to observation errors, but also that their intensities are not comparable. Where W. gives 20438·8, with intensity 4, R.R. give 20446·38 with intensity (0), and a similar effect is shown in the second line of the above triplet (R.R. have not observed so far in the red as 14813). The differences 5·6, 3·2 are comparable with the differences shown by the ν_1 sets roughly estimated above, say 5371, 5374, and would seem to be due to the same effect, although this would not go with the explanation there suggested. Further R.R.'s 20446·38 is 5633·4±2 ahead of 14813 and is therefore one exact determined value of the F_2 separation. We are justified therefore in taking 20438 as a D_{11} line belonging to the $S(\infty) = 50400+\xi$. With $d = 29964{\cdot}20+\xi'$ the calculated D line is therefore $20435{\cdot}80+\xi-\xi' = 20438{\cdot}8+{\cdot}4p$

$$\xi = 3{\cdot}0+{\cdot}4p+\xi'$$

where ξ' is small. Hence $S(\infty) = 50403{\cdot}00+\xi$ where ξ is small and equal to $\xi' \pm {\cdot}4$. A reliable value of this limit has thus been obtained. The measure of its reliability is that of two assumptions (1) that the 5649 (or 34) series studied at the beginning of the discussion of this element is an F type analogous to the 1864 of X, and (2) that 20438* is a D_{11} line. With this, $S(\infty) = 50403{\cdot}00+\xi$ and $\nu_1 = 5371+d\nu_1$, $\nu_2 = 2649+d\nu_2$ the denominators and values of Δ_1, Δ_2, calculated directly are 1·475114, 1·402290, 1·370130 and

$$\begin{aligned}\Delta_1 &= 72824-2{\cdot}06\xi+12{\cdot}571d\nu_1 &&= 40\tfrac{3}{4}\{1787{\cdot}09-{\cdot}050\xi+{\cdot}31d\nu_1\}\\ \Delta_2 &= 32160-\ {\cdot}84\,(\xi+d\nu_1)+11{\cdot}726d\nu_2 &&= 18\ \{1786{\cdot}66-{\cdot}046\,(\xi+d\nu_1)+{\cdot}651d\nu_2\}.\end{aligned}$$

The δ as determined from the Δ_2 has always been the same as those found from the D and F mantissæ, here $1787{\cdot}015-{\cdot}062\,(\xi-{\cdot}4p)$. Hence $d\nu_2 = {\cdot}53$ correct to the first decimal place and $\nu_2 = 2649{\cdot}53+d\nu_2$ where $d\nu_2 < {\cdot}1$ and $\Delta_2 = 32166$. If the δ from Δ_1 is the same $d\nu_1 = -{\cdot}3$ and $\nu_1 = 5370{\cdot}7$, $\Delta_1 = 72820$. Wherever it is different it has always been slightly less than that from Δ_2, so that we may regard 5370·7 as a maximum estimate for ν_1. It may be a few decimals smaller, but we have no direct means from observed lines to determine it more closely.

A similar treatment with the four equally probable values of δ are appended. Also changes in $d\nu_1$, if δ from ν_1 is the same as from ν_2. This will at least give maximum values of ν_1. The cases are given for the limit used above and also in the last two columns for that if 20446 is the D_{11} line, that is the limit about 7 larger.

	$d\nu_1$	$d\nu_2$	$d\nu_1$	$d\nu_2$
$1787{\cdot}87-{\cdot}062\xi$	2·5	1·7	3·5	2·5
1787·01−	− ·3	·5	·9	1
1786·13−	−3	− ·8	−2	− ·3
1785·26	−6	−2·1	−5	−2

---

* Possibly not so certain. 20446 gives one of the normal 5633 to 14818. Its weak intensity may be due, as in the other gases, to the presence of displacements and so may be D_{11}. It would make $S(\infty) = 50410$.

The limit 50403 and $\delta = 1787{\cdot}01$ require the smallest changes and seems preferable, although limit 50410 and $\delta = 1786$ is perhaps possible. Both 1787·87 and 1785·26 would appear excluded. The numerical work to follow will be based on the limit 50403, $\nu_1 = 5370{\cdot}7 + d\nu_1$, $\nu_2 = 2649{\cdot}5 + d\nu_2$, $\mathrm{S}(\infty) = 50403{\cdot}00 + \xi$. These make $\Delta_1 = 72820$, $\Delta_2 = 32166$.

With these more exact values of $\mathrm{S}(\infty)$ and ν_1 it is now possible to obtain the important constants, the *p*-links. They are found to be

$$a = 4630{\cdot}72 - 4{\cdot}11x + {\cdot}010\xi + {\cdot}74d\nu_1 \qquad c = 6277{\cdot}21 - 6{\cdot}16x - {\cdot}020\xi + 1{\cdot}347d\nu_1$$

$$b = 5370{\cdot}70 - 5{\cdot}01x \qquad d = 7399{\cdot}60 - 7{\cdot}67x - {\cdot}046\xi + 1{\cdot}822d\nu_1$$

$$e = 23678{\cdot}42 - 22{\cdot}96x - {\cdot}056\xi + 4{\cdot}912d\nu_1$$

where the terms in x denotes the changes for a displacement of $x\delta_1$ in $\mathrm{S}(\infty)$.

To determine the u, v links requires a knowledge of $\mathrm{S}_1(1)$. It is impossible to obtain really definite information on this point from direct observations. In default of this the following considerations will give some indications, and will serve to illustrate how the laws of relationship already determined can give clues indirectly. In the cases of A, Kr, X the lines for $\mathrm{S}_1(1)$ are respectively -42642, -42469, -42153. The corresponding line for RaEm should therefore be looked for about -41800, with $s(1) = \mathrm{P}(\infty)$ about $50403 + 41800$, or, say 92200. The e link unfortunately is too large to test for the lines for it would reach back to lines about 11103 **2649·5** 12751·5 **5370·7** 18122·2, of which the last only would be in the observed region. Consequently no evidence can be obtained with reference to it from the ν_1, ν_2 separations. There is a strong line at (8) 17909·9. It would give $\mathrm{S}_1(1) = 41586$, smaller than seems likely. Also it has a 5637 link not usually connected, so far as we yet know, with S lines. It however is the only strong line in the neighbourhood, and it will be probably wiser to conclude that the S (1) line has no linked e line to it. The above value for $\mathrm{P}(\infty)$ gives a v link about 13660, and this link on the supposed S (1) should produce a line about 28140. We find the set

(2) 20163·7 **2645** (4) 22808·68 [28183·7]

in which the last is extrapolated by the normal $\nu_1 + \nu_2$ from the first.

It falls in the large gap where no lines have been observed and the spectroscopic apparatus was defective. If this be really the $\mathrm{S}(1) - v$ set, it is possible to calculate what the u, v links are. Thus $\mathrm{S}_1(1) = -(28183{\cdot}7 + v)$, $s(1) = \mathrm{P}(\infty) = 78586{\cdot}7 + v$, and v is produced by the $-\Delta_1$ displacement in the denominator of this. As v by a rough determination is in the neighbourhood of 13660, it can be put $= 13660 + x$ where x is not very large and can be determined by the condition that $13660 + x$ is the change produced in $92246+x$ by a decrease of $\Delta_1 = 72820$ in its denominator. The result is that

$$\begin{aligned} \mathrm{P}(\infty) = \mathrm{S}(1) &= 92266 = \mathrm{N}/\{1{\cdot}090264\}^2 \\ u &= 11191{\cdot}8 \\ v &= 13680{\cdot}0 \\ \mathrm{S}_1(1) &= -41864{\cdot}3 \end{aligned}$$

The set of S (1) lines would then be

$$-41864{\cdot}3, \qquad -36493{\cdot}6 \text{ or } 89{\cdot}7,^{*} \qquad -33844{\cdot}1.$$

On the two latter u, $e-u$, $e-v$ links as sounders should give the following lines, below which are placed the nearest observed,

$$-u\left\{\begin{array}{ll} \ 25301{\cdot}8 \text{ or } 297{\cdot}3 & \ 22652{\cdot}3 \\ (2)\ 25299{\cdot}1 \pm {\cdot}7\ (\text{W.}) & (4)\ 22636 \pm 6\ (\text{C.R.}) \end{array}\right.$$

$$-(e-u)\left\{\begin{array}{ll} \ 24007{\cdot}1 \text{ or } 3{\cdot}2 & \ 21357{\cdot}6 \\ (20)\ 23993{\cdot}82\ (\text{R.})\ 4{\cdot}45\ (\text{W.}) & (10)\ 21357{\cdot}43\ (\text{R.})\ 6{\cdot}57\ (\text{W.}) \end{array}\right.$$

$$-(e-v)\left\{\begin{array}{ll} 26495{\cdot}3 & \ 23846 \\ \ldots & (2)\ 23841{\cdot}9\ (\text{W.}) \end{array}\right.$$

They are not satisfactory sounded lines, but are given as material at disposal only, $u.\mathrm{S}_2$ and $e.\mathrm{S}_3.u$ are really good.

There is a striking triplet which has all the aspect of being a S (1) set except that of position. It is

(2) 18611·7 (W.) **2648·9** (5) 21260·21 (W.)
(0) ... 09·94 (R.) **2652·3** (3) 21262·25 (R.) **5371·3** (10) 26633·64 (R.)

Its separations give support to the values obtained above. It is a parallel set to our adopted S (1), separated from it by 15230·7. I have not been able to recognise any arrangement of links which give this value although $2b+a$, and $e-v+b$ are close to it.†

---

* Sounded from observed line.

† There is the possibility to be kept in mind that it may be a real and independent S set of lines, not analogous to the set considered in the previous elements. If so, it must have the same limit and $s(1) = 77036{\cdot}64 = \mathrm{N}/\{1{\cdot}193177\}^2$. Now the s sequence depends on the atomic volume, although the exact relation is not known. In the next group, the alkalies, the denominator is of the form $\cdot987+kv$, where v is a number proportional to the atomic volume. If here the group constant be $\cdot99$, the denominators of the two types of s sequences can be expressed as $\cdot99+kv$ and $\cdot99+2kv$. In other words, this new type may depend on twice the atomic volume.

$S_1(2)$ should be a strong line in the neighbourhood of 25900, with $S_2(2)$, $S_3(2)$ in the ultra-violet. There are three strong lines in this region, viz. (8) 25262·73, (10) 25769·9, (10) 26633·64. The last is that which has been seen to be a parallel to $S_1(1)$. The first line gives a mantissa less than that for $S_1(1)$, and therefore would make the α constant positive, in opposition to all experience for S series. We are left therefore with 25769·9 ± 3, but this is a C.R. copper line and so far dubious. If, however, it be taken as $S_1(2)$ the formula for the S series is

$$n = 50403 - \mathrm{N} \Big/ \left\{m + \cdot129855 - \frac{\cdot039596}{m}\right\}^2.$$

The value of $S_1(3)$ calculated from this is 39112·08. Sounding with $-v$ gives 25432·08. WATSON gives a line at (0) 25432·3 ± 1 and this is corroborative so far as it goes, but the formula is out of step with those for A, Kr, X, and does not give confidence. If, however, lower intensities are admitted it is possible to obtain lines which fall in excellently with all the conditions, and moreover indicate other displaced sets. We find the lines, to which extrapolated ν_1 lines are added

(3) 25416·89 (R.)	ν_1	[30788·59]	(0) 25432·35 (W.)	ν_1	[30803·05]
(2) 25425·51 (W.)	ν_1	[30796·21]	(0) 25457·0 (W.)	ν_1	[30827·7]
			(1) 25453·1 (R.)		

If the extrapolated lines are sounded for by $-v$, the 30803·05 should give a line at 17123·05. This was seen by W. at (0) 17123·0, the nearest to this being (1) 17150·93 (R.) which might possibly be within error limits of v.[30827·7]. If, with this indication of 25432·35, 30803·0 as part of a triplet set suitable for S(2) we calculate the formula with the previously alloted $S_1(1)$ and $S(\infty)$, it comes to

$$x = 50403 - \mathrm{N} \Big/ \left\{m + \cdot101230 - \frac{\cdot010966}{m}\right\}^2$$

which is in close analogy with that for the other elements. For $m = 3$ it gives 38972·54. A v-sounder requires a line at 25292·54. ROYDS has observed (3) 25292·12, and again in this neighbourhood we find some close lines observed by only one of W. or R., viz., (8) 25262·73 (W. and R.),* (2) 25299·1 (W.) again evidencing the presence of displacements. For $m = 3$ the oun produces a change in the sequent of 3·30, so that 25299·1 which is 6·66 above $S_1(3)$ is $S_1(3)(2\delta_1)$ exactly, and 25262·73 which is 29·39 behind is $S_1(3)(-9\delta_1)$. The lines calculated from the formula for $m = 3 \ldots 8$ are given in the accompanying list, which also gives the values as sounded from an observed line. Thus in $m = 3$, the observed is entered as $v.(3)$ 38972·12, whereas the actually observed is v less.

---

* R. gives intensity 8, W. gives 3, although as a rule W.'s intensities are higher than R.'s—again pointing to changed displacements with changed conditions of excitation.

m.	Calculated.	Observed.	$d\lambda$.
1		$-\{(2)\ 20163{\cdot}7\ (\mathrm{W.})+\nu_1+\nu_2\}.v$	*
2		(0) 25432·35 (W.)	*
3	38972·54	$v.(3)$ 38972·12 (R.)	·02
4	43873·80	$e.(4)$ 43874·02 (W.)	−·01
5	46184·76	see Note	·07
6	47454·96		
7	48227·14	$u.v.(1)$ 48227·00 (W.)	·00
8	48731·22	see Note	

Note.—For $m = 3$ the e sounder is too large, and for $m = 4$ the v is too small to act.

The agreement is so good that the chosen low intensity lines for $S_1(1)$ are justified. The low intensity may be explained by the supposition that the energy of the normal line has been partly transferred to other displaced ones, in the way indicated above for $m = 3$ and as shown also by higher orders. For $m = 4$ together with 20195—the line sounded from—W. only observes (2) 20163·7, and R. only (0) 20133·0. These differ successively from the first by 30±. For $m = 4$ the oun displacement produces 1·42 on the sequent and the normal 30·5 on the limit. They are therefore possibly $e.(\delta_1)\,S_1(4)$ and $e.(2\delta_1)\,S_1(4)$. In $m = 5$ the e sounder requires 22506, not observed by either, but there is an isolated group, 22516·9 by both, (5) 22535·5 (W.), (3) 22540·4 (R.), of which 22535 is 30 ahead of the required line and 22540 is 35, so that the former is $(-\delta_1)$ on the limit. The latter, however, sounds to 46218·71 for $(-\delta_1)\,S_1(5)$ and gives $S_1(5) = 46183{\cdot}25$. The $d\lambda$ = ·07 of the list is based on this. For $m = 6$ the e requires 23776·65 with observed (0) 23760·9 (W.) and (10) 23783·75. In $m = 8$ the $u+v$ link = 24871·8 requires 23859·5. The only observed are (2) 23841·9 (W.) and (4) 23871·3 (W.) or (5) 70·00 (R.), about equally displaced on either side of the normal line.

Some space has been devoted to the consideration of the S series as on it depends the determination of the important u, v links as well as the limit for the D series. The latter is perhaps in general the more important as it affords with the F, the means of obtaining accurate determinations of the oun. In the present case, however, the measures have such large possible observation errors that they do not add to the accuracy already found in the foregoing discussion. A few points only will therefore be here referred to. Both the D satellites for $m = 1$ and the F lines for $m = 2$ with related lines form the majority of the lines in the long wave end of the spectrum ($n < 20000$). One clue as to a distinction between the two sets may be found in the fact that where the 5640 link occurs on a D line it links backwards, while in the F it links forward. Consequently where a line has a link forward it is probably not a D line, but is either a F line, or is attached to a D line (as in the 1864 set in X attached to the D group near 20312). The D-qualification test with D (∞) = S (∞) can be applied to these suspected D lines, taking the 20438 as a definite D_{11} line. That is the mantissa difference of the lines from that of 20438 must be an oun multiple. A very large number are found to satisfy this test. It must be remembered that as the oun is so large as 447 there can be no doubt as to the satisfaction of this test or not, even for the largest allowable observation errors, nor to the actual oun multiple when it is satisfied; on the contrary the possible observation errors are so great, that the observations do not enable us to increase the accuracy of the oun itself as already found. By combining a large number of cases it would no doubt be possible to diminish the probable error of its value, but the heavy work would not be justified at present, especially as there should be hope of better measures in the immediate future. At the same time the existence of the D lines may be illustrated by the following two sets for $m = 1$.

			18357 (C.R.)		(1) 21036·25
			5671 ± 4		**5633·0**
$20\frac{1}{4}\delta$	(2) 18641·4 (W.)	**5387·12**	(1) 24028·52 (W.)	**2640·71**	(1) 26669·23
			19583 (Ram.)		
			5630		
	(1) 19842·4 (W.)	**5371·3**	(2) 25213·7 (W.)		
34δ	(2) 16062 (W.)		(1) 21456·56 (R.)		
	5650		**5633·21**		
	(6) 21712·09 (R.)	**5377·68**	(2) 27089·77 (R.)		

			(0) 16445·61		(0) 19086·1 (W.)
			5652·4		**5662**
	(3) 16725·1 (W.)	**5372·9**	(4) 22098 (C.R.)	**2650·6**	(2) 24748·6 (W.)
$18\frac{1}{4}\delta$	(1) 17493·05 (R.)	**5375·2**	(1) 22868·07 (W.)		
			(1) 18357 (C.R.)		
			5636		
	(0) 18609·94 (R.)	**5383·9**	(10) 23993·82 (R.)		
$31\frac{1}{4}\delta$	(0) 14813 (W.)				
	5625·8				
	(4) 20438·8 (W.)				

In these it should be noted:—

(1) The satellite separations are in the usual ratio 5 : 3, for in the first $20\frac{1}{4} \times 5 = 101{\cdot}25$, $34 \times 3 = 102$ or for D_{23} $20\frac{1}{2} \times 5 = 102{\cdot}5$, $33\frac{3}{4} \times 3 = 101{\cdot}25$. In the second $18\frac{1}{4} \times 5 = 91{\cdot}25$, $31\frac{1}{4} \times 3 = 93{\cdot}75$. Both sets have the same undisplaced limit = 50403, and the two corresponding D_{11} lines are displaced $23\frac{1}{2}\delta$ in their sequences.

(2) The 21712 which acts as a D_{11} line shows the modified ν_1 separation to a second line but of much smaller intensity, in close analogy with what has been seen in X.

(3) The mantissa of $16725{\cdot}1 \pm {\cdot}5$ is $804601 - 26{\cdot}8\xi + 13{\cdot}4p = 25\ (32166{\cdot}16 - 1{\cdot}07\xi + {\cdot}53p) + \delta_1 = 25\Delta_2 + \delta_1$. The second line of the triplet 22098 is one of the copper

C.R. lines. Behind this is a line (2) 22079·34 W. ($d\lambda = \pm{\cdot}05$) corresponding to the $-\delta_1$ displacement in the sequent—*i.e.*, the modified D separation. With S (∞) its mantissa $= 804161-26{\cdot}77\xi-26{\cdot}77d\nu_1+5{\cdot}35p$

$$= 25\,(32166{\cdot}44-1{\cdot}071\xi-1{\cdot}07d\nu_1+{\cdot}21p).$$

If this be combined with the mantissa of $29964 = 913165 = 511\delta$ giving $\Delta_2 = 32166{\cdot}27-1{\cdot}116\xi$, there results the equation

$${\cdot}14+{\cdot}045\xi+{\cdot}21p-1{\cdot}07d\nu_1 = 0.$$

This can be satisfied by $\xi = 0$, $d\nu_1 = 0$, and p a fraction. It does not therefore help to a closer determination. With good measures it should be practicable to find $d\nu_1$ within ·05 and p a small fraction. This equation would then give the small correction for ξ and so increase considerably the degree of accuracy of Δ_2 and δ. The particular point however gained is that here is found one of the fundamental d sequences depending on pure multiples of Δ_2.

(4) They all show -5640 links when this link lands in the observed region except 22868. Where the measures are reliable they congregate round a value 5633 ±. The 22079 of the last paragraph is 5633·73 above the 16445. This is a further justification of 22079 belonging to a $-\delta_1$ displaced sequent.

With 20438 as D_1 (1), RYDBERG's table gives D (2) as in the neighbourhood 37486 ±100. A u sounder gives the region 23806. There is a line (10) 23783·75 (W.) which if linked in this way gives $D_{11}(2) = 37463{\cdot}75$. The two lines $m = 1, 2$, and S (∞) give the formula

$$n = 50403-\mathrm{N}\Big/\left\{m+{\cdot}909601+\frac{{\cdot}003564}{m}\right\}^2$$

with D (3) = 43242·02. The e-sounder requires 19563. The only line in the neighbourhood is $\lambda = 5105$ by RAMSAY, who says his measurement is very rough. If we allow $d\lambda = 5$A, the wave-number is 19583 ± 20, and it *may* be the line sought for. There seem also other D_{11} groups as in X. One instance is adduced in the next paragraph.

I end the discussion of the RaEm spectrum by a consideration of the source of the 5640 separation. In X we found the conditions satisfied by oun displacements on the F (∞) of $5\Delta_2-\delta_1$, and $2\Delta_2+6\delta_1$. But here the values of the separations themselves seem very indeterminate. The values as arranged on p. 431 would seem to point to 5633, 5649 with ν_2 about 2800 and 2820. The limit of the particular F series on which our whole discussion of RaEm is based is 29964·20. As a fact, however, this limit can only generate in the proper neighbourhood separations of 5646, 2806, and the displacements are $5\Delta_2-6\delta_1$, $2\Delta_2+2\delta_1$. The dependence on these displacements so analogous to those in X is evidence that the links in question (5640, 1864) are of analogous type. We have already seen in X that these separations occur as links, as well as direct displacements on sequences present in the wave-number. Let us determine the F limit required to give a separation in the correct neighbourhood with a displacement $5\Delta_2$. The conditions are that not only must the $5\Delta_2$ give the ν_1 but since it must belong to a $d(1)$ sequent, the mantissa of the limit itself shall also be an oun multiple. The result is that the limit should be within close limits of 29617 and that then the ν_1 will be close to 5648 one of the most probable values found on p. 431. But this $29617 = \mathrm{F}(\infty) = d(1)$ should belong to a D line $50403-27917 = 20786$. Now there is an observed line (0) 20784·31 (W.) with another at (7) 20750·34 (W.) or (7) ...52·70 (R.) behind it. Taking 20784·31 its sequent is 29618·69 and mantissa 924298 or 11133 above that of 29964·20. Now $6\frac{1}{4}\delta = 11169$ and the difference 36 requires a change in wave-number $dn = 1{\cdot}10$ or $d\lambda = -{\cdot}25$A. The line was only read to ·1A so that this error is permissible. The F (∞) is then 29617·57. The $-5\Delta_2$ displacement on this gives 35266·16 or a separation 5648·59, and on this it will be found that an extra $2\Delta_2+5\delta_1$ gives $\nu_2 = 2810{\cdot}97 = 2811$ say. Further a change of δ_1 on $5\Delta_2$ or $2\Delta_2+5\delta_1$ produces changes in ν_1, ν_2 respectively of 17·88 and 20·00. This accounts for the concomitant value of 5630·71, and 2831 also observed is due to $5\Delta_1-\delta_1$, and $2\Delta_2+6\delta_1$ precisely those in X.

The Value of the Oun.—The data for the evaluation of the oun are :—

(1) The triplet separations and S (∞). The ν_1, ν_2 are not determinable directly from S lines and consequently no exact values can be obtained from observation. The general argument on p. 433 is in favour of 5371, 2649, or about 3 larger according to limit. These are strongly supported by the set of lines on p. 437, clearly a linked S (1) or an independent S (1) set showing also the displaced S_2, S_3 by W. and R. respectively. These values point to $\delta = 1787{\cdot}0$.

(2) From the F and **F** series giving definite $\mathrm{F}(\infty) = 29964{\cdot}20$ and $f(1)$; p. 430. The definite result is a limitation of δ to one of six alternative values, of which four between 1785·26 ... 87·87 are equally probable, the ambiguity being due to imperfect measurement. This argument is independent of the values of ν_1, ν_2.

(3) The mantissa of the D_{13} line 16725 is $25\Delta_2+\delta_1$ and that of the D_{23} line 22079·34 is $25\Delta_2$. This gives

$$\Delta_2 = 32166{\cdot}44-1{\cdot}07\xi-1{\cdot}07d\nu_1+{\cdot}21p,$$

$$\delta = 1787{\cdot}024-{\cdot}059\xi-{\cdot}059d\nu_1+{\cdot}012p.$$

Here probably

$$\xi<{\cdot}5, \qquad d\nu_1<{\cdot}1, \qquad p<{\cdot}5, \text{ or, say,}$$

$$\delta = 1787{\cdot}024\pm{\cdot}05.$$

These three determinations are all independent.

Argon.—There are a large number of strong lines in the visible spectrum evidently connected with D(1) and F(2) lines. The oun is so small that the qualifying test for D lines is not so definite as in the other gases, although this is to some extent remedied by the fact that the measures on the whole are good with a possible maximum error of ·02A. We shall therefore make no attempt to discuss the F and D series with the same fullness as in the other cases. The groups selected are certainly not the only ones and possibly may not be the most important ones, but they will be sufficient to give data for the determination of the oun to about the same degree of accuracy as for the other gases of this family. Take for the first group.

$m = 1.$

(1) 23782·51 (K.) **179·92** (6) 23962·43 (E.V.)
$71\frac{1}{2}\delta$
(2) 23899·08 (K.)

$m = 3$

(1) 44827·37) **175·51** (1) 45002·88
$39\frac{1}{4}\delta$
(3) 44835·41

$m = 2$

(3) 39357·06 **178·95** (4) 39536·01
$171\frac{3}{4}\delta$
(3) 39420·68
$39\frac{1}{4}\delta$
(5) 39439·33

The lines for $m = 2, 3$ are by EDER and VALENTA. Again the low intensity for $D_{11}(1)$ is to be noticed, but we have here some indication of the source of this peculiarity. EDER and VALENTA give the line at 23899·83 of intensity 9, and state that it appears only with a very strong condenser discharge. The difference in the two measures ($d\lambda$ = ·13A) may possibly be due to observation although it is greater than is the rule between the measures of K. and of E.V. As an oun displacement in the sequent produces a change dn = ·4, the difference may be due to $2\delta_1$ displacement. If so, no error will be introduced in the succeeding considerations by overlooking this, treating it as due to observation error and using KAYSER's measurements, subject to a smaller possible error. For in either case the dependence of the sequent mantissa on the multiple law is not affected and although the multiple itself is different, at the same time the difference in the two readings will not modify the formula constants to an extent to appreciably affect the calculated lines for $m > 2$.

Using the limit $D(\infty) = S(\infty) = 51731{\cdot}03$ and the first two D_{11} lines the formula found is

$$n = 51731{\cdot}03 - \mathrm{N}\Big/\left\{m + 989074 - \frac{{\cdot}003976}{m}\right\}^2.$$

This gives for $m = 3, 4, 5$, the lines $n = 44834·16$, 47323·04, 48672·56. The O−C for the first is $d\lambda = -·06$ whilst the others should be weak and are near to the limits of observations. The links are so small that as sounders they produce lines also close to the observed region. For $m = 1$, the difference of the sequent mantissæ of 23899 and 23782 (D_{12}) is 4144 and $4\Delta_2 - 6\delta_1 = 71\frac{1}{2}\delta = 4141·2$, or exact with $d\lambda = ·013$ on two lines. It will be shown later (p. 447) that D_{12} depends on the Δ_2 multiple. The series is, therefore, probably a real doublet series with this for the outer satellites. For $m = 2$ the D_{13} set with a displacement of about $171\frac{3}{4}\delta$ is inserted since it recalls the order of magnitude in Kr and X, whilst the two D_{12} satellites for $m = 2$ and 3 are inserted because they show the same displacement of $39\frac{1}{4}\delta$. An extrapolated line on an observed one [23611·5], **179·5**, (2) 23791·0 gives a displacement 175δ for $m = 1$ corresponding to that for $m = 2$. It may also be noted that in Argon the third lines of the triplets appear to be missing in this D series.

The following arrangement would seem to indicate the existence of a set parallel to the above :—

23899·08	**20·74**	(3) 23919·82	**188·73**	(1) 24108·55
39439·33	**20·65**	[39459·98]	**171·60**	(3) 39631·58
44835·41	**20·65**	[44856·06]	**177·22**	(4) 45033·28
[47323·01]	**20·66**	(47343·67)	**178·97**	(1) 47522·64

The [ ] in the second column are calculated from the first column by adding 20·65. The line in ( ) is determined as the mean of the linked lines *u.*(1) 47784·02 = ... 44·55 and (1) 46183·62.*u.e* = ... 42·80. As 5δ on the limit produces a separation of 20·65, the lines in the second column would correspond to the parallel displaced series (-5δ) D_{11}. The separations of the last column increase with the order to a normal value of ν_1 and point to a constant displacement in the sequent.

With the above D set should be associated a doublet F series whose $F_1(\infty) = d_{11}(1) = 51731·03 - 23899·08 = 27831·95$ and separation that of the D satellite or 116·5. In searching for this allowance will have to be made for the prevalence of displacement in the lower orders. As a fact, however, this only appears for $m = 3$ in F_1, although in F_2 there are considerable signs of it. The following table with succeeding notes contains the data, including also those for the corresponding **F** series ; also for certain lines which show a separation about 150, and which may probably be the analogue of the 1864 X series.

We shall denote this last set by dashing the letters.

The formula obtained from the two first F_1 lines and the given limit is

$$n = 27831·95 - \mathrm{N} \Big/ \left\{ m + ·757701 + \frac{·006544}{m} \right\}^2.$$

The O−C between observed (or deduced from observed) and calculated are given under the heading *d*λ. As before ( ) refers to deduced lines [ ] to calculated.

m.	F.	*d*λ.	27831·95.	F.
1	−(4) 7404·33 (P)	*		[63066]
	115·65			
	−(4) 7288·68 (P)			...
	...			...
2	(1) 13444·52 (P)	*	*31·42*	(42218·33)
	116·12			**115·89**
	(1) 13560·64 (P)			(42334·22)
	150·91			
	(1) 13595·43 (P)			...
3	[20073·40]	?	*31·14*	(35588·89)
	116·85			**114·3±2**
	(20190·25)			(35703·2±2)
	151·05			
	(4) 20224·41			...
4	[89·50]			
	(10) 22991·57	−·39	*30·13*	(32668·69)
				116·31
	...			(32785·0)
				150·50
	...			(2) 32819·19
5	[24·49]		[*30·88*]	
	(5) 24521·94	·42	*29·60*	(1) 31137·27
	115·31			**116·57**
	(24637·25)			(31253·84)
				150·29
	...			(31287·56)
6	(2) 25429·28	·17	32·12	(1) 30234·96
	116·54			**116·85**
	(25545·82)			(4) 30351·81
	152·84			
	(1) 25582·12			...
7	(26010·63)	−·19	33·29	(3) 29655·76
	116·44			
	(26127·07)			...
	...			...
8	(4) 26402·01	−·07	30·99	(29259·97)
	114·27			**118·82**
	(26516·28)			(29378·79)
	147·75			
	(5) 26549·76			...
9	(1) 26679·62	−·01	34·96	(3) 28990·30
				119·00
	...			(29109·30)
	153·75			**150·51**
	(1) 26833·37			(1) 29140·81

m.	F.	$d\lambda$.	27831·95.	F.
10	(5) 26885·78	− ·29	34·10	(1) 28782·42
	...			... **148·23**
	...			(1) 28930·65

Notes.—Displacements are in evidence except for F_1. A desired wave-length can be obtained very closely since a δ_1 displacement produces a change of only ·42. Consequently in deductions by displacement the results have little weight, and in fact would have none at all were it not that we now know as a fact that such exist. Failing such likely displacements, recourse has been had in a few cases to linked lines and here their evidence has weight.

$m = 2$. For F_1, (-16δ) (1) 42192·13 = 18·21; (9δ) (1) 42233·12 = 18·45; mean taken. For F_2 (-5δ) (1) 42326·07 = 34·22.

$m = 3$. There is no observed F_1, but (3) 19351·20.*e* = 20070·91 or with E.V. 71·65;

F_1 (3) $(7\Delta_2)$ = 20103·85 and 20105·52 is observed by E.V. $d\lambda = -·42$.

For F_2, $(-38\delta_1)$ (4) 20174·83 (E.V.) = 90·22; $(20\delta_1)$ (2) 20198·39 = 90·29 mean taken.

For F_1, (-5δ) (1) 35580·73 = 88·88, also F_2, (-5δ) (1) 35695·03 E.V. ± 2 = 703·18 ± 2.

$m = 4$. The observed line for F_1 is much stronger than should be expected; also its O − C is in the opposite direction to that of the others. It may be displaced by Δ_2 on the sequent and so intensified. The Δ_2 would produce O − C = ·00 and the observed = F_1 (4) (Δ_2), or it may hide the real F_1. F_1 is inserted as 32819 − 150·50. It only indicates that 32819 is the correct value of F′. F_2 is given as (5δ) (1) 32793·11.

$m = 5$. F_1 shows a link e = 719·81 to (3) 25241·75. F_2 is entered as (18δ) (4) 24665·59 = 37·25, but has no weight. $F_1 = (15\frac{1}{2}\delta)$ (1) 31279·10.

$m = 6$. F_2 is the mean of $(-6\delta_1)$ (2) 25536·15 = 45·87 and (7δ) (1) 25557·12 = 45·78. The calculated value of F_1 gives a better 151·7 to F'_1.

$m = 7$. The calculated F_1 is 26009·30. The deduced value is (1) 25290·92.*e* = 26010·73, which again has an apparent approximate e link = 720·77 forward to an E.V. line at (2) 26731·40. This is however a coincidence as the last line is S_2 (2). In connection with the linked line 26010·73 may be taken the pair of lines (-6δ) (3) 26001·11 = 10·89 and $(7\frac{1}{2}\delta)$ (3) 26022·85 = 10·63 whose mean agrees precisely with the former. For F_2 a similar split with two may be observed $(-2\frac{1}{2}\delta)$ (3) 26123·03 = 27·07 and (2δ) (1) 26130·61 = 27·30.

$m = 8$. F_2 is (5) 25796·57.*e*; also (-2δ) (2) 26512·75 = 16·05. It may be noticed that the separations to F_1 and F′ are both about 2·3 too small, or $1\frac{1}{2}\delta$ on the limit. F_1 is (2) 28540·26.*e* and F_2 is (1) 28659·08.*e*. These linked lines in this order are therefore reliable.

$m = 9$. F_2 is (1) 28389·59.*e*.

It should especially be noted how with increasing order the normal calculated lines appear as observed with good intensities. As in the other instances it suggests itself that this is due to the diffusion of the energy of the lower order lines into the formation of numerous displaced ones. The usual accurate determination of the limit as the value of $\frac{1}{2}(F_1+F_1)$ is not here applicable as, for the reason given above, the determination of the actual displaced lines is unreliable. These values are printed in italics. The more reliable results point to a limit higher than that calculated in S (∞), with ξ about +2.

The Value of the Oun.—The preceding results for the D and F series afford material for the more exact determination of the oun. The D_{11} line 23899 must have a mantissa which is a multiple of the oun, though not necessarily of Δ_2. The mantissa of the $F(1) = -7404$ should be a multiple of Δ_2. PASCHEN's estimate for the maximum error of the F(1) is 1A, and KAYSER's for the D(1) is ·02A. These mantissæ are, introducing the actual errors as $d\lambda = p_1$, $\cdot 02\, p_2$

$$985098+4\cdot 06p_2-35\cdot 66\xi$$

$$764245+13\cdot 7p_1-25\cdot 03\xi = 723\,\{1057\cdot 047+\cdot 0190p_1-\cdot 0346\xi\}.$$

In the case of F (1), given that its mantissa is a multiple of Δ_2 the multiple must be 723 for 724 or 725 would alter Δ_2 by 1·5, whereas we already have its value as close to 1057·0. The mantissa of d_{11} is so large that by itself it is not possible to determine the oun multiple. But the difference of the mantissæ of d_{11} and f is much smaller and also an oun multiple. It is

$$220853-13\cdot 7p_1+4\cdot 06p_2-10\cdot 66\xi$$
$$= 209\,\{1057\cdot 047+\cdot 0190p_1-\cdot 0346\xi\}-5\delta_1+2-17\cdot 7p_1+4\cdot 06p_2-3\cdot 67\xi.$$

Even this is too large to settle the multiple on account of the observation errors and uncertainty in ξ. It may be an oun more or less. The following consideration, however, will give some indication on this point. The mantissa of the satellite 23782 as found above was $4\Delta_2-6\delta_1$ above that of D_{11}. Consequently its complete mantissa is $209\Delta_2-5\delta_1\pm\delta_1-(4\Delta_2-6\delta_1) = 205\Delta_2+\delta_1\pm\delta_1$ above that of $f(1)$ or $= 928\Delta_2$ or $+\delta_1$ or $+2\delta_1$. Now this is so close to the Δ_2 multiple as to point to the fact that this satellite is the fundamental one, in which the rule is a multiple of Δ_2 for the mantissa when there is no relative displacement with the second or third of the triplet satellite set. But here the observed ν_1 is $179\cdot 92 = 179\cdot 50+\cdot 42$ and ·41 is the change produced by an oun displacement. In other words the mantissæ of the D_{12}, D_{22} lines are either $M(\Delta_2)$, $M(\Delta_2)+\delta_1$, or $M(\Delta_2)-\delta_1$, $M(\Delta_2)$. But it cannot be $M(\Delta_2)-\delta_1$. Hence if the multiple law holds here it must be in the D_{12} satellite, the $-\delta_1$ must be taken above, or the difference is $209\Delta_2-6\delta_1$ and $16\cdot 4-17\cdot 7p_1+4\cdot 06p_2-3\cdot 69\xi = 0$. This is easily satisfied by moderate values of the p's and ξ. Further, the mantissa of 23782 is $928\Delta_2$.*

This makes the mantissa of $D_{11} = 932\Delta_2-6\delta_1$, whence

$$932\Delta_2 = 985098+6\delta_1+\ldots = 985184\cdot 9+4p_2-35\cdot 66\xi$$
$$\Delta_2 = 1057\cdot 064+\cdot 004p_2-\cdot 0382\xi.$$

---

* It may be noted in passing that the satellite separation is 116·57, but is reproduced in the F series as 115·65 (P.) for $m = 1$, 116·12 (R.) for $m = 2$. The latter, more reliable, is ·45 less than the corresponding satellite separation. This general effect is therefore completely explained by the F limit for the second series being d_{22} in place of d_{12}.

The D_1 satellite gives

$$928\Delta_2 = 980956+p_3-35{\cdot}44\xi, \quad \text{obs. error} = {\cdot}02p_3$$
$$\Delta_2 = 1057{\cdot}064+{\cdot}04p_3-0382\xi.$$

Also the $f(1)$ gives

$$\Delta_2 = 1057{\cdot}047+{\cdot}0190p_1-{\cdot}0346\xi.$$

The value of ξ in $f(1)$ is not precisely the same as in the others as it involves the observed error in D_{11}. If ξ is the value for $d_{11}(1)$ that for $f(1)$ is $\xi-{\cdot}12p_2\,(d\lambda = {\cdot}02p_2)$ and

$$\Delta_2 = 1057{\cdot}047+{\cdot}0190p_1+{\cdot}004p_2-{\cdot}0346\xi.$$

Thus p_1 is of the order ·9 and the best value is

$$\Delta_2 = 1057{\cdot}064-{\cdot}0382\xi\pm{\cdot}004.$$

The value of ξ is probably a very small positive quantity. The O$-$C for D_{11} (4) is $dn = 1{\cdot}25$ pointing to a positive ξ. The data for the accurate determination from $\frac{1}{2}(F+\mathbf{F})$ are defective. Of directly observed lines those for $m = 6$ appear reliable as the separations are good. This gives the limit as 27832·12 or $\xi = {\cdot}17$. Failing any better determination we put $\xi = {\cdot}17+\xi$

$$\Delta_2 = 1057{\cdot}057-{\cdot}0382\xi\pm{\cdot}004. \qquad \delta = 57{\cdot}9209-{\cdot}0021\xi\pm{\cdot}0002.$$

Probably

$$\xi \not> 1 \qquad \text{and} \qquad \delta = 57{\cdot}921\pm{\cdot}002.$$

The manner in which so many independent conditions fit in to give this value is very remarkable, and should give great weight to the accuracy of the above value. It must not, however, by itself be regarded as a definitive value free from all possible doubt without further support from independent groups of D or F series. This is not easy to obtain. In the first place the spectrum of A seems to approximate more to the doublet sets of He, and we lose the advantage of dealing with triplets. Again, it is clear from the way in which the lines are crowded in the region below $n \doteq 2300$ and in which modified separations in the neighbourhood of 180 preponderate that a large number of displaced groups must exist, and displaced groups have generally shown a fragmentary quality. It is difficult, however, to deal with these in the same way as in Kr, X, or RaEm, because the oun is so small that the separations produced by one oun displacement are comparable with observation errors or uncertainties in the limits.*

---

* I had prepared the details of an additional D group with corresponding F and **F** series depending on a displaced D(∞). As however, while supporting the above value of the oun, it does not decrease its possible error, no present object is to be gained by printing it.

The separation 150 has been referred to as probably the analogue of the 1864 F separation in X. The reasons for this supposition are based on its magnitude and its occurrency curve. This separation in Kr, X and RaEm has been explicable as due to a displacement close to $5\Delta_2$ on the mantissa of a *d* (1) sequent. In the present case the example taken is not the most important F series, but its limit will be near that to which the separation is due. The displacement required in it to produce a separation 150 is very close to $5\Delta_2$, in fact $5\Delta_2$ produces 149. The separation in question then is caused by displacement of the normal Δ_2-multiple for this series. The second reason is based on the form of the occurrency curve, which shows the same sharply defined single peaked curve as in X. It is represented in Plate 2, fig. 5.

It may be interesting to note that all STARK's A^{+++} lines, *i.e.*, whose sources have lost three electrons, all show *e* links except one. They are

			(3) 23674	**719·23**	(3) 24393	
			(3) 23696			
			(2) 24053	**719·22**	(4) 24772	
721·90	(1) 23639	**719·46**	(7) 24359			
		? F_1 (5).	(5) 24521	**719·81**	(3) 25241	**720·24**.

Neon.—The principal sources* for measurements in the spectrum of Neon are observations by LIVEING and DEWAR, BALY, and WATSON. These have been supplemented by interferential measures by PRIEST, MEGGERS, and MEISSNER. ROSSI has added a few lines down to 2352. Through the kindness of Mr. W. F. MEGGERS I have also had the advantage of using an as yet unpublished list of very accurately measured lines made by himself and Messrs. BURNS and MERRIL at the Bureau of Standards in Washington.† The lines by LIVEING and DEWAR are only roughly measured, but as in the case of the other rare gases comprise many not observed by others. BALY's list extends from 6717 to 3037, WATSON's from 7245 to 2736 and contains a considerably larger number of lines. Both in BALY's and WATSON's lists considerable gaps appear with only a few strong lines, especially between 4250 and 3500. These are filled by a number of weak lines observed by LIVEING and DEWAR. These latter are very important for the complete discussion of the Ne spectrum as they represent the scattered displacements and linked lines of low order series lines which normally should occur as single strong lines but which are here wanting.

---

* LIVEING and DEWAR, 'Roy. Soc. Proc.,' vol. 67, p. 467 (1900); E. C. C. BALY, 'Phil. Trans.,' A, vol. 202, p. 183 (1903); H. E. WATSON, 'Roy. Soc. Proc.,' A, vol. 81, p. 181 (1908); J. G. PRIEST, 'Bull. Bur. Standards' (U.S.A.), vol. 8, p. 2; W. F. MEGGERS, 'Bull. Bur. Standards,' vol. 12, p. 198 (1915); K. W. MEISSNER, 'Ann. d. Phys.,' vol 51, p. 115 (1916); R. ROSSI, 'Phil. Mag.,' vol. 26, p. 981, (1913).

† Referred to below as B.M.M.

Unfortunately L.D.'s measures are only given to 1A, and thereby their value is greatly diminished as they become merely indicatory and cannot serve as quantitative data. The accuracy of **Baly** and of **Watson** is good and probably about the same. **Priest** claims an accuracy with probable error $<$·0005A, **Meissner** with error not $>$·0015, but the accuracy of an interferometer measure depends very largely on the nature of the individual line. **Meggers'** results are exceptionally valuable in that he gives interferometer measures of a number of lines of small wave-length 3701 to 3370 where the S (2) and some of the higher order F lines occur. **Rossi** has succeeded in allotting lines to series.

Neon affords an apparent exception to the rule amongst the rare gases of different spectra, according as they are developed with or without condenser in the tube discharge. On the other hand its spectrum would appear to be a composite one of the typical "red" and "blue" spectra. It undoubtedly has a portion analogous to the "blue" as will be seen by the results obtained below, completely analogous to those found in this communication for the other gases, which refer to their "blue" spectra. On the other hand in some remarkable sets of accurately equal separations discovered by **Watson*** it shows a relation to the analogous well-known constant separations observed by **Rydberg** in Argon. Further, it is specially rich in lines in the red region. In the list of lines observed at the Bureau of Standards referred to above there are 225 between 8783 and 5689. Since in each periodic group of elements the number of lines as a whole increases very rapidly with the atomic weight, the excess of red lines in Ne is even comparatively greater than the actual number shows. The majority of these lines are weak, but they almost all fall into a few definite linkages in which the links are the constant separations discovered by **Watson**. Some of these special linkages again are connected together by the p and s links, especially the *e.u.v.* They belong to the F type of order $m = 2$, and should afford most valuable information as to the way in which parallel and displaced lines are related. I hope to return to this question on a later occasion, and only refer to them in the present discussion incidentally as they afford some evidence for the determination of the value of the oun.

The wave-numbers of observed lines published stretch from 13251 to 36536. From analogy with the spectra of the other gases we must not therefore expect to find more than one order in each of the S and D series. Nor, with its small atomic weight will the *e.u.v* links be large enough to act as efficient sounders. On the other hand the whole of any F series ($m = 1$ excepted) should lie within the above limits. It is therefore clear that the attack on the problem must be made first on this series. One datum at least is at our disposal in the magnitude of the oun. Taking the atomic weight at 20·0 ± ·01 the calculated value of δ is 14·47 ± ·01. This value of 14·47 may therefore be treated as exact to one or two units in the last digit.

---

* 'Proc. Camb. Phil. Soc.,' vol. 16, p. 130 (1911); 'Astro. Journ.,' vol. 33, p. 399 (1911).

The F (2) line should be a strong one in the neighbourhood of 17000. There are a number in this region. The lines in the following list are found to give a good series and is doubtless *one* F series. In this case we possess the great advantage of very accurate interferometer measure of the 1st, 3rd, and 4th lines in international units. These are used to determine the formulæ constants. The wave-lengths are given as measured, the wave-numbers are all in Rowland units.

NeF.

m.	λ.	*n.*	O.	O−C.
2	(10) 5852·48 I.	17081·46	·00	*
3	(0) 4157·74 R.	24044·88	·05	−·05
4	(5) 3700·89 I.	27009·47	·00 ?	*
5	(5) 3501·22 I.	28552·37	·00	*
6	3393 R.	29464·2	1·00	−·66
	(3417.*e*) R.	(29454·2)	1·00	(·47)
7	3329 R.	30030·57	1·00	·58

Notes to Table.—For the first line PRIEST gives λ = 5852·4862, MEISSNER ·4875, and MEGGERS ·488. These all give the same wave-number to the second decimal place. The second is a weak line by WATSON not observed by BALY. BALY gives a line, intensity 4, *n* = 24039·45 not observed by WATSON. We have here a concrete observational example of the facility with which a normal F line of low order can split up into displaced lines by slightly different excitations. In this case the mantissæ difference in the sequents is 159 and $11\delta = 159·1$, so that 24039 is $F(3)(-11\delta)$. The third line is a strong line observed by both WATSON and BALY, but the measure used is deduced from an interferentially measured line *n* = 28439·801 by deducting the WATSON link separation 1429·429 (both accurate). WATSON'S measure is 27009·95. The line for $m = 5$ is by MEGGERS, but WATSON gives the same *n*. The remainder of the series comes in one of the gaps referred to above in which only L.D. have observed. They give lines which may serve for $m = 6, 7$. Also for $m = 6$ there appears a linked line at $29257·2+e = 29454·2$ (using the value of *e* found below). The calculated wave-numbers for $m = 8, 9, 10$ are 30422·6, 30703·3, 30906·5. No lines are observed between 30203 and 30722. The last two have lines by L.D. near them at 30722·7 ± 9, 30922·3 ± 9.

The formula as found from $m = 2, 4, 5$ is

$$n = 31850·19 - N \Big/ \left\{m + ·794726 - \frac{·139254}{m}\right\}^2.$$

The calculated value for F (1) is λ = 12241·97 *in vacuo.*

There can be little doubt about these lines forming a series, but there is so far no independent evidence that it is of the F type beyond its analogous position in the spectrum to that of the other rare gases. If it is of this type we should expect to find a number of parallel series as well as of the associated **F** type. It is then necessary to test for these conditions.

A considerable number of parallel series can be arranged all weak for $m = 3$ as in the original series. Some of these are given in the following list in which the numbers in clarendon below the wave-numbers give the separations from the corresponding F lines. Those to the right of the vertical line are added. The others are deducted.

(1) 16768 { **2·67** (0)...71 **312·73** { **5·56** (0)...74	(8) 16816 **265·36**	(2) 16831 **249·59**	(2) 16845 **235·84**	(4) 16889 **191·63**
23623 **315·2 ± 3**	—	(0) 23798 **246·69**	(1) 23812 **232·58**	(2) 23854 **191·29**
26695 **314·6 ± 3**	*26745* **264 ± 3**	*26766* **244 ± 3**	(1) 26776 **233·33**	*26816* **193 ± 6**
	u.(9) 28396 **262·83**	(1) 28302 **250·28**	*28216.u* **228 ± 6**	—
	(3) 29196‡ **261·7**	—	—	*29257* **200 ± 9**

(5) 16925 **155·86**	(1) 16948 **113·00**	(5) 16996 **85·23**	(5) 17222→ **141·51**	(7) 17342 **261·22**
—	—	—	*24183* **138 ± 6**	*24312* **267 ± 6**
26860 **150 ± 6**	(26892)* **117 ±**	(2) 26922 **87·10**	(4) 27148→ **138·83**	*27285* **275 ± 6**
(9) 28396 **156·33**	(5) 28438 **113·61**	(4) 28474 **87·57**	*e*. (5) 28888 **138·68**	(28819)† **267 ± 6**
(3) 29196.*u*‡ **155·0**	*29343* **115 ± 9**	*29369* **89 ± 9**	(1) 29592 **134·15**	*29727* **268 ± 9**
			30175 **139 ± 9**	*30194.u* **265 ± 9**

The wave-numbers in italics are by L.D. and subject to probable errors $d\lambda = \cdot 5$ and possible $d\lambda = 1$. For values of links see below.

* *u.26997·6* = 26891·0; *266948.e* = 26891·8.

† *u.28927* = 28820·3; *28711.u* = 28817·7, mean taken.

‡ One at least must be a coincidence, *or* the two series can simply be linked by *u* or *v*.

Still more striking, and as will be seen later, important, are parallel series with the separations discovered by WATSON. In addition are found also others at 1932 behind F (2). They are

(5) 15149·24	**1932·22**	F (2)	**1429·38**	(6) 18510·84	**422·33**	(1) 18933·34
22103	**1941·7 ± 5**	F (3)		see Note*		
25087	**1922·9 ± 6**	F (4)	**1429·43**	(6) 28438·85	**417·44**	(4) 28856·29
(3) 26628·53	**1924 ±**	F (5)	**1426 ± 8**	*29976·5*	**424**	*30203.e*†

* F (3) has a link 423 to *24467* ± 5 and then 1432 to *25899* ± 5 suggesting a mesh in which the required ine is wanting. Here the 1932 separation goes better with the strong displaced F (3) 24039·45, giving 1936 ± 5. The mesh should be

F (3),	24044	**423 ± 6**	*24467*	**1432 ±**	*25899*
		1429·42	[25474·30]	**425 ± 6**	

† Note that 1426 + 424 = 1429·4 + 420·6 ± 9.

Any lines of the **F** type up to $m = 4$ will unfortunately lie in the ultra-violet beyond the observed region. **F** (4) should be 36690, and the largest frequency observed by WATSON is (1) 36536. The others should be weak and in a region where glass apparatus would only allow strong lines to be registered. **F** (5) should be at 35148·01 + 2ξ but is not seen. The line 35259·2 is about a v link ahead, in fact v.35259·2 = 35152·4. It may, however, be noticed that 36536 above is just 154 behind the expected **F** (4), so that it is the **F** line corresponding to the parallel F set above with the separation 156 (say F′). In the same series is also found **F′** (6) = (2) 34087·1 corresponding to the F′ (6) = 29196.u. These are of value in that it gives the means of determining the limit with great exactness. Denote the parallel series by F′. For $m = 2$ using B.M.M.'s measure for F′ (2) the separation is 17081·46 ± 0 − 16925·43 ± ·05 = 156·03 ± ·05. For $m = 5$ both lines have been measured interferentially and the separation is 28553·342 − 28397·167 = 156·175, correct to the second decimal place. The two separations differ by more than the allowable observation error, and is possibly due to the common change in sequent for series with different limits. In these cases in the separation with the larger m, this effect is very small. Consequently we are justified in taking the separation as 156·17 ± 0. For $m = 4$ we have F (4) and **F′** (4) but only a L.D. line for F′ (4). The separation, however, gives its exact value as 27009·47 − 156·17 = 26853·30. **F′** (4) is 36536·62 ± ·66 ($d\lambda$ = ·05). The mean gives the limit for the F′ series as 31694·96 ± ·33 and consequently for F as 31851·13 ± ·33, *i.e.*, ξ = ·94 ± ·33.

But further in the neighbourhood of calculated **F** (6) = [34242] are found also (1) 34336·06, (3) 33918·08 respectively 94 ahead and 323·9 behind it. In analogy also are found (5) 17176·34, (3) 16757·91 respectively about the same amounts ahead of and behind F (2), but no other corresponding F (m) lines appear. We are justified in taking the first two lines as really in the relation indicated, and are thus enabled to arrive at the value of F (6). Using B.M.M. measures the separations given by the F (2) lines are their differences from 17081·46 or 94·88, 323·55 with very small errors $\pm$·05. These, therefore, give for **F** (6) 34336·06$\pm$·25$-$94·88 = 34221·18$\pm$·25 and 33918·08$\pm$·5+323·55 = 34221·63. Both are therefore within error limits of their mean 34221·40. With this for **F** (6), and the limit corrected as above to 31851·13 the value of F (6) is 29460·96. With the value calculated from the formula O$-$C = $-$·23.

As illustrating the way in which D and F sequents are subject to displacements, it is interesting to notice that although the separations 94, 323 do not appear as directly dependent on F lines after $m = 2$, they nevertheless occur in the neighbourhood. It will be sufficient here only to refer in detail to the case of F (4) = 27009·47. At about 50 ahead of this there is a line (6) 27060·60 (W.). With this there is the following scheme :—

26766·3 (L.D.) **96·6** 26863·4 (L.D.) **197·2** (0) 27060·60 (W.) **321·4** 27382 (L.D.)

As the L.D. lines are subject, even if correct to the nearest A.U., to equally probable errors between $d\lambda = \pm$·5, or here to $dn = \pm 4$, these separations correspond to the 94·88, and 323·55, whilst 197·2 is the link *e*. The displacement of 27060 from F (4) may then be in the F sequent or the limit—probably the former, or *a*, *b* or *c* link.

The existence of the parallel sets, the indications of the associated **F** types, and the nature of the displacements all point to the conclusion that our original series is of the F type. Consequently the limit F (∞) = 31851·13 is a d(1) sequent, but there is nothing as yet to show whether it belongs to a d_{11} or a satellite set. If, however, we can find a D (1) set the value of D (∞) can be obtained with sufficient accuracy to obtain the values of Δ_1, Δ_2 and the *e* link. The further consideration of the F series will therefore be postponed until this further information has been obtained.

It has already been remarked that only the S (2) and D (1) lines in any S or D group can be expected within the observed region. A superficial inspection of the list of lines shows a very large number of separations in the neighbourhood of 46 to 49 and about 20, clearly related to ν_1, ν_2 values as they show the normal ratio $\nu_1 : \nu_2$, and are in step with those of the other gases. In about the region in which the D lines should be expected BALY gives the strong set (4) 21200·90 **49·24** (4) 21250·14 with an equally strong line at (4) 21230·11 which might be a D_{11} line to the doublet D_{12} set. WATSON, however, gives other strong lines as well, including a triplet. These give

(1) 21156·77 (W.) **48·18** (1) 21204·95 (W.) **19·71** (4) 21224·66 (B., W.)
44·13 **45·19**
(4) 21200·90 (B.) **49·24** (4) 21250·14 (B.)
(5) ·81 (W.) **49·36** (5) ·19 (W.)
29·21
(5) 21230·11 (B., W.)

We are justified by its form in taking this as a D (1) set for a preliminary trial, although the satellite separations are not in the usual ratio of 5 : 3. It is not so clear that the particular F (∞) just obtained is a d sequent belonging to it, nor, if so, whether it is a d_{11} or a satellite sequent. The latter point, however, will have little effect for our immediate object—the attainment $\Delta_1 . \Delta_2 . e$—as the differences are small and may be included in an undetermined ξ. The probability is that as this seems the only prominent D triplet set, it belongs to the normal group and that our F (∞) belongs to it, although on this point something will have to be said later. We shall take on trial that $F(\infty) = d_{11}$. In this case

$$D_1(\infty) = 21230·11 + 31851·13 = 53081·24 + \xi$$

This should also be $S_1(\infty)$ and should give the separations by oun multiples in the denominator of the sequent. The three S or D limits would then be

$$53081·24 + \xi \quad \mathbf{49·24} \quad 53130·48 + d\nu_1 + \xi \quad \mathbf{19·71} \quad 53150·19 + d\nu_1 + d\nu_2 + \xi$$

where ξ *may* be considerable, owing to uncertainty as to 31851 being of d_{11} or satellite type. The mantissæ of these are

$$437428 - 13·540\xi, \qquad 436752 - 13·521\xi, \qquad 436485 - 13·514\xi$$

which give as differences

$$\Delta_1 = 666 - ·019\xi + 13·5 d\nu_1 = 46\,(14·477 + ·29 d\nu_1 - ·0004\xi)$$
$$\Delta_2 = 267 - ·007\,(\xi + d\nu_1) + 13·5 d\nu_2 = 18\tfrac{1}{2}\,(14·432 + ·73 d\nu_2 - ·0004\xi),$$

in which it must be remembered that calculations with seven-figure logarithms give uncertainties of a unit in the last digit. The direct calculation of δ from the atomic weight has already given $\delta = 14·47 \pm ·01$. This gives $46\delta = 665·6 \pm ·46$, $18\tfrac{1}{2}\delta = 267·69 \pm ·18$. Thus these limits give without any doubt the true oun multiples in Δ_1, Δ_2, and the calculated Δ_2 then gives a closer value of $267·7 \pm ·2$. No possible change in ξ can affect these results. Further, these multiples are quite in step with the march in the other gases. The remarkably close agreement sustains the allocation of the F (∞) to the d sequence of this set, although not necessarily to d_{11}.

As a further test the satellite separations should be due to oun displacements in the sequences. These separations are 29·21 and 45·19, taking the latter because 21224 is a good measure (B. and W. agree) and the observed $\nu_2 = 19·71$ agrees so closely with the Δ_2 value and the ν_1 is subject to the very common triplet modification. These separations require displacements in the sequence mantissæ of $851 - ·040\xi + 29 d\nu$,* and $1313 - ·060\xi + 29 d\nu$. Now $59\delta = 853·7 \pm 1·2$, $90\tfrac{3}{4}\delta = 1313·1 \pm 1·8$. This is sufficient to give the satellite multiples, but as the oun is so small, $\delta_1 = 3·62$, the close agreement cannot serve as evidence one way or the other as to the satellite nature of the doublet and triplet lines in question. The points against the allocation of these lines to the normal D group are, (1) the satellite separations are not in the usual ratio 5 : 3 (they are in fact close to the ratio 3 : 5), (2) there is no appearance of satellite lines corresponding to the parallel F lines noted above, and (3) the limit 53081 is somewhat larger than we should expect from the march of the limits in the other gases.

---

* WATSON's value of the separation 29·30 makes this 2·6 larger = 853·6.

Taking it, however, as the limit, it is possible with the given value of Δ_1 to calculate the link *e*. The result is $197+4d\nu_1$. The occurrency curve as found from B.'s and W.'s observations is given in Plate 2, fig. 6. The dotted line is the result when the numerous rough measures by L.D. are included. As is seen the maxima occur at 195·6 and 198, the calculated, at a minimum. The peaks look as if analogous to corresponding peaks found in other elements, but the analogy is doubtful. In the other elements these (much larger ones) are produced by the prevalence of displacements by a few ouns, and by Δ_1 operating on $(x\delta)$ S (∞). Such changes here would be very small and the corresponding effects are really shown by the flattened tops of the peaks. The peaks themselves are due in all probability to another cause to be considered shortly (p. 458).

In the region in which the S triplets should be expected are found (all interferential measures)

(6) 28787·86 **49·81** (5) 28837·67 **18·61** (4) 28856·28

with intensities in the proper order, although the ν_1, ν_2 are slightly different from those obtained in the D set. The interferential measures of MEGGERS and of B.M.M. differ considerably (·008) and are not so reliable therefore as usual. BALY and WATSON agree in giving the separations as 49·72, 18·64. The third line, however, has already appeared in the quite definite relation 1429+417 ahead of F (4). It cannot be S (2), but the latter may be a weak line close to it.

There is also another doublet set

(10) 27818·89	**47·75**	(1) 27866·64 }
	49·93	(1) 68·82 }
(8) ·73	**48·07**	(1) 6·80

The suggested S_2 line is abnormally weak. There are, however, here a number of other weak lines which have the appearance of displaced debris. If the first set form a S group, the second would belong to a group with smaller limit, which is also indicated by the smaller triplet separations. The limit 53081·24 was obtained from apparently the only stable D set, whilst the first set are apparently the only stable strong S group. It is natural therefore to take its limit as the same. If so, the limit for the second, being 969·03 (or ·35 $d\nu$) less, is 52112·20. The two mantissæ are $437420-13{\cdot}54\xi$ and $450723-13{\cdot}92\xi+13{\cdot}92d\nu$. Their difference is

$$13303-{\cdot}38\xi+13{\cdot}92d\nu = 20(665{\cdot}15-{\cdot}019\xi+{\cdot}7d\nu) = 20\Delta_1.$$

Furthermore the observed separation 47·75 requires, on this new limit, a displacement $664+13·9\,d\nu_1$. With $d\nu_1 = ·11$ we have Δ_1 the same as before, *i.e.*, the 47·86 is the proper ν_1 separation of the new limit. The ν_2 proper to this limit should be 19·15 giving $S_3 = 27885·90$. This is not seen, but the *u*-linked line is possibly given by L.D.'s line $27995·5\pm4-106·8 = 27988·7\pm4$. The actual S_2 line appears split up into the additional displaced lines, each (1) 27858·26, (1) 27868·82, (1) 27873·63 from observed S_2. As a parallel *u*-link to the first we find also L.D., $27964·2\pm4-106·8 = 27857·4\pm4$ for actual 58·26. The oun is too small to decide whether the displacements are produced in the limit or the sequent.

The limit for this new set, 52112·20, is more in step with the progression of the other gases, viz. :—

Ne.	A.	Kr.	X.	RaEm.
52112	51731	51651	51025	50403

It is possible to get an estimate of the *u.v* links although there is no means apparent at present of getting the exact value of $s(1)$. Either of our S groups gives the same value for $s(2)$, viz., $53081·24-28787·86$ or $52112·20-27818·89 = 24293·3$. The denominator of this is 2·1248, so that the $\mu+\alpha/2$ of $s(2)$ is ·1248. We can get an estimate of the value of $s(1)$ from the law that $(\alpha+\frac{1}{2}\Delta_1)/(\mu+\frac{1}{2}\Delta_1)$ is about ·2 in the other periodic groups. The corresponding values in this group are A, ·189 ; Kr, ·198 ; X, ·222. If the ratio ·18 is taken for Ne $\mu+\alpha$ comes to about ·1139 with $s(1)$ about 89000. These are only rough estimates, but the values of *u.v* will only alter slowly with considerable changes in $s(1)$. The values calculated from this $89000+\xi$ with $\Delta_1 = 666$ are

$$u = 106·78+·0017\xi, \qquad v = 106·86+·0018\xi.$$

This shows that although the value of $s(1)$ may be extremely rough, those of *u.v* may be relied on within a few decimals.

. One is inclined to regard the weak S set as that which is analogous to those determined in the other gases, and which certainly belong to the blue spectrum, and that in Neon, which shows only one spectrum, it is composite, and the blue not strongly developed. The question naturally arises whether there is any evidence of an unstable or weak D set with the same limits, *i.e.*, about 969 behind the former D. There is a set in this neighbourhood

(0) 20173·82 **42·86** (2) 20216·68

of the right order of inverted intensities and a small ν_1, possibly a case of the common D triplet modification. With the limit 52112·20 the mantissa of 20173 is 853094, which is 373 greater than that of the old *d*. This is within error limits of $26\delta = 376$ This may be explained by the supposition that the *d* sequence is not affected by the limit and that the present pair run parallel to one 26δ above the old D_{13}. If this existed it should be 21142·64, 21191·98, 21211·69. There are lines for the last two but none for the first. They are

[21142·74] **[49·24]** (0) 21190·61 **20·10** (6) 21210·71.

There is thus considerable evidence for the existence of parallel S groups as well as of parallel D ones. Each set has its corresponding ν_1, ν_2 separations, but with the same oun multiples for Δ_1, Δ_2. We should consequently expect to find the presence of corresponding *a* ... *e* links. With the S limit 52112 the *e* link is found to be 191·49. Now the occurrency curve for *e* gives a maximum between 195·6 and 196 pointing to a S limit as basis about 650 less than 53081 or, say, 52430. This makes $e = 195{\cdot}63$, $\nu_1 = 48{\cdot}9$. Do we find evidence for S and D sets about this amount less in wave-number than the old ? For the S we are landed in a region which forms a gap in B.'s, or W.'s observations but which contains a number of lines by L.D. Amongst these we find

$$\begin{array}{rlcl} & S_1. & & S_2. \\ \left.\begin{array}{r}\mathbf{1429}\\ \mathbf{195{\cdot}7}\\ \mathbf{105{\cdot}9}\end{array}\right\} & 27964\pm4 & \mathbf{47} & 28011\pm4 \\ & \quad\mathbf{213} & & \\ & 28177\pm4 & \mathbf{40} & 28217\pm4\end{array}$$

The suggested S_1, S_2 fall into the proper positions, the link 1429 is one of WATSON's constants, 195·7 is the *e* link proper to the limit, 105·9 is *u* or *v*, and 213 is 2*u* or 2*v*.

For the D sets we find amongst W.'s lines

$$\begin{array}{lcl} (0)\ 20551{\cdot}35 & \mathbf{49{\cdot}62} & (2)\ 20600{\cdot}97 \\ \qquad\mathbf{157{\cdot}51} & & \\ (3)\ 20708{\cdot}86 & & \end{array}$$

The 20551 is 601 below $D_{13} = 21156$, which when the variability of the satellites is considered may be taken as the analogue of 610 for the S set. If it is corrected* by $dn = {\cdot}7$ $(d\lambda = -{\cdot}17)$ the ν_1 becomes the correct value 48·9, and it is then also separated from 20708 as a D_{11} line by 156·8, which at once suggests the origin of the 156 parallel F and **F** sets previously brought to light. If this relation be real, the old $F(\infty)$ 31851 is a *d* satellite sequent, belonging to 20551·35 or ...2·0 and the limit is $20552{\cdot}0 + 31851{\cdot}13 = 52403{\cdot}13$. Not only is the 156 separation found, but there are a large number of lines in the neighbourhood which give, possibly within the error limits, the other separations indicated by the traces of parallel F series adduced at the beginning. Also as indicating a D region we find large repetitions of *e* and *b* links characteristic of D linkages. As referring to the F series we find, starting from 20551—which corresponds to 31851—

---

* Not necessarily error, probably the usual D displacement on sequent.

					(3) 20556·84		
				44·13			
	(0) 20551·35	**49·62**	(2) 20600·97				
			(0) ... 2·02		**85·81**		
	114·84				(0) 20642·61		
	(5) 20666·19						
	157·51						
D_{11}	(3) 20708·86						
	194·42						
	(2) 20745·77						
	232·02						
	(3) 20783·37						
	249·22						
	(1) 20800·57						
		269·39			**312·66**		
			[20820·74]	[**48·78**]	(2) 20869·52		
						3·04	
			275·00		(0) ... 72·56		**5·61**
						2·57	
			(1) 20826·35	**48·78**	(4) ... 75·13		

There is also (1) 20542·57 at 166·29 behind the D_{11} line 20708·86 which seems related to ROSSI's series referred to later.

The application of the D qualification test is unfortunately here nugatory on account of the smallness of the oun. It is, however, striking that so many of the F separations put themselves in evidence. This agreement adds considerable weight to the allocation. The evidence is also practically convincing when the WATSON separations are brought into evidence as is done a few paragraphs below. Support is also given in the doublet system shown in connection with 20826, which show the same small displacements as are exhibited in the first of the parallel F sets adduced above. In this case to 20869 as a D_2 would correspond a [20820] as a D_1 269·39 ahead of 20551, whereas the F is 312·73 behind the corresponding F_1 line (17081). If the two sets correspond, *i.e.*, the *d* sequent and the F(∞) for this set are the same, the D group must belong not to 20551 but to a parallel D group in which the limit is 312·73−269·39 = 43·34 behind the normal D(∞). Now this is the separation of 20556 behind 20600. In fact the separation of 20556 and 20869 is 312·66. This further indicates that 20556 is a D_2 line in the position shown in the above table, in which the corresponding D_1 line is too weak to have been seen. This also gives another F separation—85—with 20642. The change 44·13 on the D limit requires a displacement on the mantissa of 611 and $42\frac{1}{4}\delta = 611·3$, which is exact. The ν_1 corresponding to this new limit, with $\Delta_1 = 666$, is 48·30. The difference between this and the observed 48·78 is just within our assumed possible errors on two lines.

Again, the double line 20600·97, 02·02 apparently reproduces itself in the F group, as MEISSNER says 17081 shows also a weak component. It is also in evidence in another relation as we shall see immediately.

The existence of parallel F series with WATSON'S separations and also 1932 has already been pointed out, viz.,

$$\bullet\ \mathbf{1932}\quad \mathrm{F}(1)\quad \mathbf{1429}\ \bullet\ \mathbf{422}\ \bullet$$

If these depend like the others on D series, the same separations should be found in the reverse order. They do not appear with D_1, but they are seen with the stronger line $\mathrm{D}_2 = 20600$, with which also the linkage relation enters. Thus

$$\mathbf{1428{\cdot}46}\quad (8)\ 18753\quad \mathbf{419{\cdot}13}\quad (1)\ 19172\left\{\begin{array}{ll} \mathbf{1428{\cdot}68}\quad (2)\ 20600{\cdot}97\left\{\begin{array}{l}\mathbf{1925\pm3}\quad \mathit{22526}\\ \\ \end{array}\right. & \\ & \left.\begin{array}{l}\mathbf{1429{\cdot}45}\\ \mathbf{1428{\cdot}40}\end{array}\right\}(4)\ 22030\\ \mathbf{1429{\cdot}73}\quad (0)\ 20602{\cdot}02 & \end{array}\right.$$

in which there is also a mesh between 19172 and 22030 and a link back from 18753. There may also be a forward **1932** linked line if L.D. has $d\lambda = -1{\cdot}1$. It is striking that a mesh is repeated in the F (1) set. Thus

$$17081{\cdot}46\quad \begin{array}{ccc}\mathbf{1429{\cdot}38} & 18510{\cdot}84 & \mathbf{422{\cdot}50}\\ \mathbf{1430{\cdot}24} & 18511{\cdot}70 & \mathbf{421{\cdot}64}\end{array}\quad 18933{\cdot}34$$

The separation of the 20600 lines is practically repeated in the 18510.

The parallel D set, with $\mathrm{D}_2 = 20556$, also shows traces of the separations, with the mid lines not seen, in analogy with the weak mid one just considered. Thus

$$(7)\ 18709\quad \mathbf{417{\cdot}86}\quad [19127]\quad \mathbf{1429{\cdot}42}\quad (3)\ 20556{\cdot}84$$

The absence of the corresponding lines connected to 20551 may be due to the scattering of the lines by displacement. The wanting lines should be [19122·07], [18704·6], the former corresponding to the weak 19172 of the second D group. Now we find strong lines corresponding to both these at the same separations ahead and suitable for the same displacements, viz.,

$$\begin{array}{rr} [19122{\cdot}07] & (1)\ 19172{\cdot}29\\ \mathbf{20{\cdot}79} & \mathbf{20{\cdot}27}\\ (5)\ 19142{\cdot}86 & (4)\ 19192{\cdot}56 \end{array}$$

These numerous F and D relations render it certain that the sets of lines adduced belong respectively to sets of F series and the D series. Moreover, it suggests that the source of the 1429, 417 separations is the *d* sequent or $F(\infty) = 31851{\cdot}13$.

A possible supposition is that their source should be in the $S(\infty)$ limit. If so we should expect it to appear strongly in the S lines, and so in the $S_1(2)$ lines considered above. As a fact, however, there is no sign of such in either of the S groups adduced, except a very dubious one 1425 between two L.D. lines each of which has an ambiguity ± 4. It takes place between $S_1 = 27964$ and 26539 backwards, so that if its source were here it would be a positive displacement on $S(\infty)$ or a negative one on $s(2)$, both unusual. The strongest argument for its source being in the 31851 is that the separations in question show themselves in *all orders* of $F(m)$—in other words, occur in the limit $F(\infty)$.

The Value of the Oun.—It has already been found that the value of the oun calculated from the chemist's atomic weight is $14{\cdot}47 \pm {\cdot}01$ and that the oun multiple for Δ_2 is $18\frac{1}{2}$ or $\Delta_2 = 267{\cdot}70 \pm {\cdot}02$. This is too small—or the inexactness too large—to obtain a more accurate value as in the other cases directly from the F or D mantissæ. It is, however, possible to arrive at an extremely accurate estimate by proceeding step by step with successive approximations, and for this purpose the F separations are clearly at disposal. The wave-lengths of many of the F (1) lines are very accurately known (B.M.M. will be used), they are so large that the *dn* are small multiples of $d\lambda$, and being of order $m = 1$, an oun displacement will produce a comparatively large change in *n*. In spite, therefore, of the smallness of the oun it is possible to get some definite information. The reliability of the information will depend on two assumptions—

(1) That the lines employed are F lines parallel to the series $F(1) = 17081$.
(2) That no displacements occur in the *f* sequents themselves.

If the assumption (2) is not satisfied the series in question will not show constant separations from the corresponding $F_1(m)$ lines, but will converge or diverge with increasing order. The lines we shall make use of have been measured probably up to a few thousandths of an Ångstrom, and the accuracy is greater than one in the seventh digit in the value of *n*. Moreover, in calculating with seven-figure logarithms, in which also we have to do with differences between two numbers, errors amounting to unity or more are liable to enter. Consequently where these very accurate numbers occur nine-figure logarithms have been used. As the wave-lengths are given in I.A. the calculations have been made on that basis. The limit $31851{\cdot}1300$ R $= 31852{\cdot}1816$ I. $N = 109678{\cdot}6$. Put $\Delta_2 = 267{\cdot}70 + x$, therefore $\delta = 14{\cdot}4703 + {\cdot}054x$ where at present x lies between $\pm{\cdot}2$. It should be noted that in the *d* sequences, the satellite displacements are not in general multiples of δ, but of $\delta_1 = \frac{1}{4}\delta$. The correct value of a wave-length will be taken as the observer's value $-{\cdot}001 \times p$, so that $dn = + \ldots$.

Amongst the sets of F series given (p. 452) two, in addition to F (1) itself have been measured to the required degree of accuracy. They are those showing separations of 85 and 265. Under these conditions

$$\text{Mantissa of } 31852{\cdot}1816+\xi = 855630{\cdot}30-29{\cdot}130\xi$$

$$\text{Wave-numbers of } 5852{\cdot}4870 = 17082{\cdot}0220-{\cdot}0029p_1$$

$$\mathbf{85{\cdot}4080-{\cdot}0029\,(p_1-p_2)}$$

$$5881{\cdot}8958 = 16996{\cdot}6140-{\cdot}0029p_2$$

$$\mathbf{265{\cdot}3518-{\cdot}0029\,(p_1-p_3)}$$

$$5944{\cdot}8344 = 16816{\cdot}6702-{\cdot}0029p_3$$

(1) *Separation* = 85·41.—This is the same within error limits as occurs in the D series, but the corresponding F series shows ν increasing to 87·57 W or 87·41 B at $m = 5$, which means additional displacements either in F (∞) or the sequences. As the 85 agrees in both the F (1) and D series this will not happen in the F (∞), and the separation 85·41 will be due only to the actual separation in 31851. The limit of the F series in question is therefore $31852{\cdot}4170-85{\cdot}4080 = 31765{\cdot}8336+\xi$.

$$\text{Its mantissa} = 858133{\cdot}15-{\cdot}085\,(p_1-p_2)-29{\cdot}247\xi$$

$$\text{Difference from } F_1 = 2492{\cdot}85\ldots$$

Now

$$9\Delta_2+5\tfrac{3}{4}\,\delta = 2492{\cdot}504+9{\cdot}311x, \qquad x < {\cdot}2$$

Hence

$$9{\cdot}311x = {\cdot}35-{\cdot}085\,(p_1-p_2)-{\cdot}117\xi$$

$$x = {\cdot}038-{\cdot}009\,(p_1-p_2)-{\cdot}0126\xi$$

The important point to notice is that with our preliminary limit of uncertainty ($x < {\cdot}2$), the oun multiple cannot be any other than that used, so that the second approximation is quite definite. It has already been seen in the discussion of the F_1 series that ξ is probably within ±·33 also p_1-p_2 will not numerically be greater than 4. Hence $x = {\cdot}038 \pm 036 \pm 0042 = {\cdot}038 \pm {\cdot}04$

$$\Delta_2 = 267{\cdot}738 \pm {\cdot}04$$

$$\delta = 14{\cdot}4705 \pm {\cdot}0009$$

(2) *Separation* = 265·3518.—Limit = $31586{\cdot}8298+\xi+{\cdot}0029\,(p_1-p_3)$

$$\text{Mantissa} = 863408{\cdot}31-{\cdot}085\,(p_1-p_3)-29{\cdot}498\xi$$

$$\text{Difference from } F_1 = 7778{\cdot}01-{\cdot}085\,(p_1-p_3)-{\cdot}368\xi$$

Now

$$29\Delta_2+\delta = 7777{\cdot}77+29{\cdot}05x$$

therefore

$$29{\cdot}05x = {\cdot}23-{\cdot}085\,(p_1-p_3)-{\cdot}368\xi$$

$$x = {\cdot}008-{\cdot}0029\,(p_1-p_3)-{\cdot}0126\xi = {\cdot}008 \pm {\cdot}016$$

$$\Delta_2 = 267{\cdot}708 \pm 016$$

In comparing these values it must be remembered that ξ and p_1 enter in both. Equating the two, ξ disappears and

$$\cdot 030-\cdot 006p_1+009p_2-\cdot 003p_3=0$$

This is easily satisfied by possible observation errors—*e.g.*, $p_1=-p_2=p_3=1\cdot 4$ say,

$$\Delta_2=267\cdot 713-\cdot 0126\xi\pm\cdot 01$$

(3) WATSON'S *Separations.*—When the strong lines giving these separations are taken, the exactness of the equality of the separations obtained from them is most remarkable. Using the interferentially measured lines by PRIEST, MEISSNER, and MEGGERS, with 9-figure logarithms the following values are found in I.A. for the means

1429·4292	8	·0065	·0048
417·4533	7	·0120	·0064
1070·075	1		

The last enters as a component of 1429, viz., 1429 = 1070+359.

The second column gives the number of the observations used, the third the maximum deviation of a single observation from the mean, and the last the root mean square of all the deviations. Using these the mantissæ of the following numbers have to be found with $\xi+d\nu$ added—

31852·1816 **1429·4290** 33281·6106 **417·453** 33700·0636, **1070·075** 32922·2530

They are

$$\begin{array}{rl}
855630\cdot 30-29\cdot 130\xi & \\
& \mathbf{40286\cdot 52-1\cdot 857\xi+27d\nu_1} \\
815343\cdot 78-27\cdot 273(\xi+d\nu_1) & \\
& \mathbf{11278\cdot 99-\cdot 505\xi+27d\nu_2} \\
804064\cdot 79-26\cdot 768(\xi+d\nu_1+d\nu_2) & \\
825224\cdot 33-27\cdot 721(\xi+d\nu_3) & \mathbf{9980\cdot 55-\cdot 448\xi-27(d\nu_3-d\nu_1)}
\end{array}$$

In this particular case, $F_1=17082$, and F_1+1429 have both been observed interferentially and the separation is $17082\cdot 0240-18511\cdot 4499=1429\cdot 4259$, or $d\nu_1=-\cdot 0031$. The 417 separation is altered by some displacement to $422\cdot 52\pm\cdot 17$ and is therefore not directly applicable. The calculations have been carried out on the basis of the values obtained on the averages. It must be remembered, however, that when displacements enter in a sequent they frequently occur on values of the sequent which have already received a small displacement, in which case the separations themselves receive small changes. Too much weight must therefore not be given to the 417 case here, in which its actual value for the particular set is not obtainable. The $d\nu_2$, however, is certainly very small.

The discussion of the two F separations has given $\Delta_2 = 267{\cdot}713 - {\cdot}0126\xi \pm {\cdot}01$, $\delta = 14{\cdot}4710 - {\cdot}00068\xi \pm {\cdot}0005$. Let x denote the correction required on this value of Δ_2. Then

$$36\Delta_2 + 16\tfrac{3}{4}\delta = 9880{\cdot}057 - {\cdot}4653\xi + 36{\cdot}90x$$
$$42\Delta_2 + 2\tfrac{1}{2}\delta = 11280{\cdot}123 - {\cdot}5306\xi + 42{\cdot}13x$$
$$150\Delta_2 + 9\delta = 40287{\cdot}19 - 1{\cdot}896\xi + 150{\cdot}48x$$

Supposing that the true multiples of the oun are y greater, and putting $d\nu_1 = -{\cdot}0031 + d\nu_1$

$$36{\cdot}90x - {\cdot}017\xi + 27\,(d\nu_3 - d\nu_1) - {\cdot}40 + 3{\cdot}62y_3 = 0$$
$$42{\cdot}13x - {\cdot}025\xi - 27d\nu_2 + 1{\cdot}13 + 3{\cdot}62y_2 = 0$$
$$150{\cdot}48x - {\cdot}040\xi - 27d\nu_1 + {\cdot}75 + 3{\cdot}62y_1 = 0$$

or

$$x = {\cdot}010 + {\cdot}00046\xi - {\cdot}73d\nu_3 - {\cdot}097y_3$$
$$x = -{\cdot}027 + {\cdot}00057\xi + {\cdot}64d\nu_2 - {\cdot}086y_2$$
$$x = -{\cdot}0050 + {\cdot}00026\xi + {\cdot}18d\nu_1 - {\cdot}024y_1$$

In these ξ cannot be more than a few units, $d\nu < {\cdot}02$ and $x < {\cdot}01$. This can only happen if all the $y = 0$. Thus again there is the very important fact that the oun multiples are quite definite and are those used in the actual calculations. ξ is not large enough to affect the limits of accuracy in x. The separation 1070 is not so well determined as the others and $d\nu_3$ may well be $> {\cdot}01$. Thus the first and third can easily give the same values of x, but the second would require $d\nu_2$ of the order ·03, inadmissible if the ν_2 were accurately determined. But as a fact the average ν_2 as we have seen does not enter in the line here considered and it may be so large as to alter the multiple. The second may therefore be considered as not at disposal, and the third then gives very close limits, viz., with $d\nu_1 \not> {\cdot}01$

$$\Delta_2 = 267{\cdot}708 - {\cdot}0124\xi \pm {\cdot}002$$
$$\delta = 14{\cdot}4708 - {\cdot}0007\xi \pm {\cdot}0001$$

the same value, though with closer limits of accuracy, as was obtained from the 265 separation. With maximum $\xi = {\cdot}33$, $\delta = 14{\cdot}4708 + {\cdot}0003$.

But further, in addition to WATSON's separations, we have found affixed to the F series, another $= -1932$, and this must be tested. The linked line is given by B.M.M. as 6598·953 I.A. Still using 9-figure logarithims, the wave-number is 15149·7338 giving the separation 1932·2902, and corresponding to a limit $31852{\cdot}1816 - 1932{\cdot}2902 = 29919{\cdot}8914 + \xi - d\nu$. The mantissa of this is $914613{\cdot}17 - 31{\cdot}997\ (\xi - d\nu)$. The displacement on 31852 is therefore $58982{\cdot}87 - 2{\cdot}867\xi + 32d\nu$. With the above values of Δ_2, δ, $220\Delta_2 + 6\delta = 58982{\cdot}58 - 2{\cdot}728\xi \pm {\cdot}44$. This again is an exact agreement.

The foregoing does not give Δ_2 with the desired definiteness unless the value of ξ is determined. The reason is that it has been based on displacements on the same limit.

An independent datum is necessary to fix the value of ξ, which at present we know cannot be greater than a few units. In the other gases this is obtainable by the conditions that the mantissa of $f(1)$ is a multiple of Δ_2 and that of F (∞) of δ_1. The value of F (1) is known with great exactness, probably to within less than ·001A. **Priest** gives $\lambda = 5852{\cdot}4862$, **Meissner** ·4875, B.M.M. ·488. They all agree within ·001 of ·487 so that the wave-number of F (1) is $17082{\cdot}0242+{\cdot}0029p$ with p within ± 1. But here we have to answer the question whether our first F line 17082 is really the first of the series. If this F series—a very marked and definite one—belongs to the 1864XF type it has its first line for $m = 1$ in the ultra-red (calculated $n = -8168{\cdot}62\lambda = 12242$) and far beyond the reach of any sounders. This uncertainty will therefore have to be kept in mind. The mantissa of 17082 must in any case be a multiple of δ_1, and may be of Δ_2 (if it is the first F line). In addition the D set 20551 may belong to that D satellite set with M (Δ_2), or since the ν_1 separation is modified, differ from a multiple of Δ_2 by a few ouns only. We have then

$$\text{Mantissa F}(\infty) = 855630{\cdot}30-29{\cdot}130\xi = \text{M}(\delta_1) \text{ and possibly M}(\Delta_2)$$

$$\text{,,} \quad f(1) = 725012{\cdot}36-92{\cdot}257\xi-{\cdot}2675p = \text{M}(\delta_1) \text{ and possibly M}(\Delta_2).$$

The term in p will not affect our immediate purpose and may be omitted. There is evidence for $\xi < {\cdot}33$. Also $\Delta_2 = 267{\cdot}708-0124\xi+x$ with $x < {\cdot}002$. We find

$$2708\Delta_2 = 724953{\cdot}26-33{\cdot}58\xi+2708x, \qquad 2708x < 5{\cdot}4$$

$$3196\Delta_2 = 855594{\cdot}77-39{\cdot}63\xi+3196x, \qquad 3196x < 6{\cdot}4$$

Hence

$$\text{Mantissa of F}(\infty) = 3196\Delta_2+35{\cdot}5+10{\cdot}5\xi-3196x$$

$$\text{,,} \quad f \quad = 2708\Delta_2+59{\cdot}10-58{\cdot}68\xi-2708x.$$

A first definite result is that 20551 is not a D line of the M (Δ_2) type. Its mantissa must differ from such by at least 7 or 8 ouns. The coefficient of ξ is so large that the actual oun multiple cannot be uniquely decided. Further f cannot follow the Δ_2 multiple law unless ξ be of the order 1, or three times our estimated limit of variation. We cannot say that it is impossible. If however this F line is the first of the series then ξ must be of this magnitude. It is seen that no further definite and certain information can be obtained. It will however be of some interest to follow out the assumption that its mantissa is a M (Δ_2). In this case put $\xi = 1+\xi$ where now ξ is small, and

$${\cdot}42-2708x-58{\cdot}68\xi-{\cdot}2675p = 0$$

The uncertainties in ξ are now ${\cdot}42\pm 5{\cdot}4-59\xi\pm{\cdot}26 = 0$, or $\xi < {\cdot}1$.

$$\Delta_2 = 267{\cdot}6957-{\cdot}0124\xi+x.$$

In conclusion some reference is necessary to the series due to **Rossi**.* These consist of two doublet series each with separation of order 167 and a third singlet series. The two sets of doublet series appear to converge to the same limits, in the same way as the satellites of a D series. The first lines given by him are of the order $m = 3$, but amongst the B.M.M. as also the MEISSNER lines those for $m = 2$ are also seen. Further, there appear close strong companions to the second lines of doublets in the first series, and in this they recall the behaviour of the F lines in the alkaline earths* and so suggest that they are really lines of the F type, but belonging to a different *f* sequence from that discussed above. If so the limit must behave as a displaced value from the 31852. The sets are, in ROWLAND units:—

---

* *Loc. cit.*

m.					
2			**165·44**	(8) 11932·92	
					1·73
	†	(4) 11767·48	**167·17**	(0) 4·65	
		107·63			
	†	(2) 11875·11	**168·29**	(4) 12043·40	
3				(8) 17342·68	
					1·06
	†	(5) 17176·58	**167·16**	(1) 3·74	
		46·39			
	†	(5) 17222·97	**168·21**	(5) 17391·18	
4	†	(6) 19677·64	**166·58**	(6) 19844·22	
		23·92			
	†	(4) 19701·56	**167·86**	(5) 19869·42	
5	·04	(5) 21033·94	**166·87**	(5) 21200·81	
		14·08			
	−·02	(4) 21048·02	**167·59**	(3) 21215·61	
6	·04	(2) 21850·62	**167·08**	(3) 22017·70	
		10·17			
	−·25	(1) 21860·79	**169·63**	(4) 22030·42	
7	·06	(3) 22380·18	**167·18**	(2) 22547·36	
		9·91			
	−·01	(0) 22386·29‡			
8	·07	(0) 22742·78	**167·31**	(0) 22910·09	
9	·20	[23168−167]		23168 (L.D.)	

---

* [III., p. 383, *seq.*]

† Lines used in calculating formulæ.

‡ This line was not included by ROSSI.

The formulæ calculated from the first three lines of each series give

$$n = 24104{\cdot}67 - \mathrm{N}/\{m + {\cdot}973110 + {\cdot}016932/m\}^2$$
$$n = 24104{\cdot}48 - \mathrm{N}/\{m + {\cdot}987215 + {\cdot}014948/m\}^2$$

The O−C values for these are given in the above list before the wave-numbers. They are larger than we should expect the observation errors to be. If, however, the limit be reduced by ·40 and μ, α calculated from the first two, the O−C for $m = 4$ is −·04 and zero for all the others, $m = 5...8$. For $m = 9$, using L.D. 23168±3−167 as the observed for the first of the pair the O−C is ·2. The agreement for all is therefore exceedingly close except for $m = 4$. The calculated wave-number for this is 19677·50 which makes the separation 166·72 and more in step with the others. The uncertainty in the limit 24104·67 must therefore be very small.

The doublet separations show a tendency to converge with increasing order, but this is clearly due to the fact that the constant separation must be taken between the strong first line and the weak second. After $m = 3$ the weak is not seen and the observed separation is not the true one but that between the first and the second strong one. The separation in the second series is somewhat larger than in the first, that for the first being 167·17 and for the second 168·25±04 (mean for $m = 2, 3$). These require limit mantissæ changes of 7354 in the first where $308\frac{1}{4}\delta = 7354{\cdot}1$, and of 47·79±1·76 extra for the other where $3\delta = 46{\cdot}41$. The separation 1·73 between the four satellites in the first doublet must be due to displacement in the sequent. It requires a mantissa change of 209 and $14\frac{1}{2}\delta = 209{\cdot}8$.

These considerations point strongly to the conclusion that the series are of the F type. Fortunately, owing to the fact that the first line of the first series has been very accurately measured by B.M.M., as $\lambda = 8495{\cdot}380$ I.A., it is possible to test if the mantissa of $f(2)$ is M (Δ_2). Taking the observation error as $-{\cdot}001p$, the wave-number is $11767{\cdot}8680 + {\cdot}0014p$. Its mantissa with limit $24105{\cdot}0731 + \xi$ (the I measure of 4·2700 R.) is

$$981622{\cdot}48 - 120{\cdot}87\xi + {\cdot}17p = 3667\,\{267{\cdot}6910 - {\cdot}03296\xi + {\cdot}00005p\} = 3667\Delta_2$$

within the uncertainty of ξ.

The first line of the second series has not been measured so exactly. Its $\lambda = 8418{\cdot}38 - {\cdot}02p$; $n = 11875{\cdot}50 + {\cdot}028p$; mantissa $= 13092 - 1{\cdot}56\xi + 3{\cdot}52p$ larger than the other. Now $49\Delta_2 - 7\delta_1 = 13091{\cdot}8$. Hence if the two limits are the same 24104·89 (*i.e.*, put $\xi = -{\cdot}19$), the mantissa is larger by $49\Delta_2 - 7\delta_1 + {\cdot}51 + 3{\cdot}42p$. It is satisfied by $p = -{\cdot}14$ or $d\lambda = {\cdot}002$. But it is possible that these F are due to independent groups, *i.e.*, that the sequence of the second series also depends on a whole multiple of Δ_2. This cannot happen unless the two limits are different, which in fact seems to be the general rule. If the limit is displaced by $y\delta_1$, ξ is $3{\cdot}617y/44{\cdot}25 = {\cdot}082y$, and the mantissa difference from $49\Delta_2$ is now

$-7\delta_1+{\cdot}51+10{\cdot}04y+3{\cdot}42p = 0$ or $y = 2{\cdot}4-{\cdot}34p$. In other words $y = 2$, $p = 1$. This makes the limit of the second series less than that of the first by ·17, in agreement with the relative values as actually found by direct calculation from the observed lines. Although therefore the limits suggest equality, there is good evidence that the observed difference, though small, is real, and is due to a displacement of two ouns. The sequence depends on an origin of $3667+49 = 3716\Delta_2$.

ROSSI's third series is (6) 28788·08, (2) 34087·11, (1) 36536·62, (1) 37869·0, the last being observed by himself. The first three give

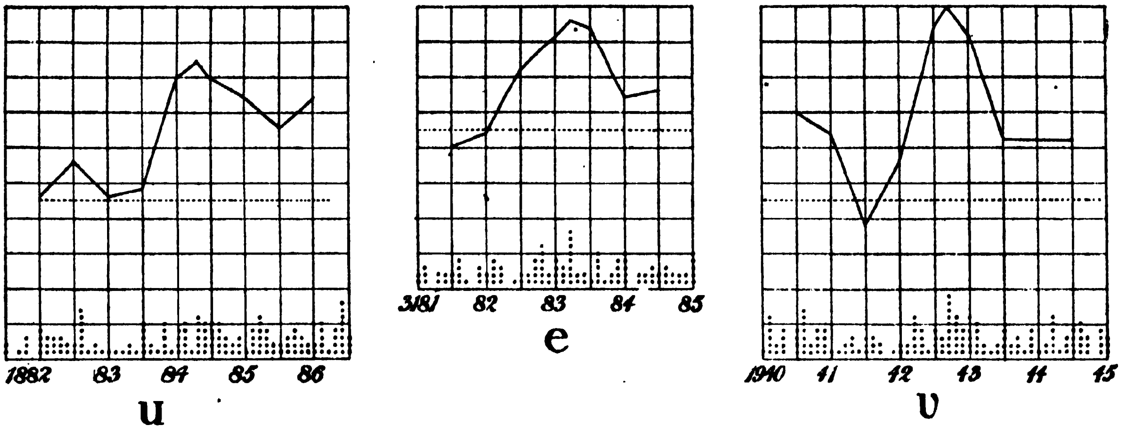
$$n = 40896{\cdot}73-\mathrm{N}/\{m+{\cdot}024118-{\cdot}043614/m\}^2$$

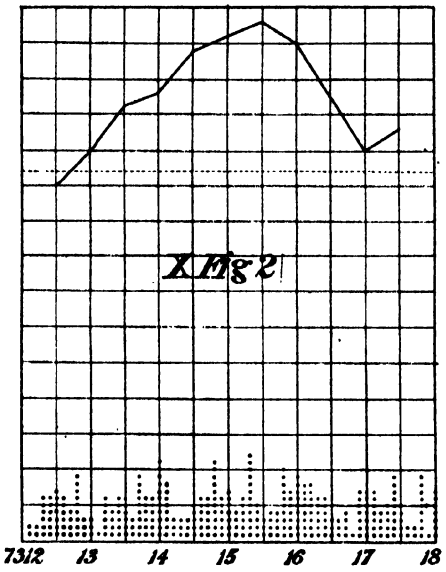
with $\mathrm{O}-\mathrm{C} = -{\cdot}36$ for the last, probably excessive. The first line is the line adopted as one of the S_1 (2) group above, the third is F (4). I feel some doubt, therefore, as to these forming a real series, although the sequence has all the appearance of the *s* (*m*) type for these gases. The limit does not seem to have any relation to the doublet set. Its denominator is about 1·63.

TABLE of Constants.

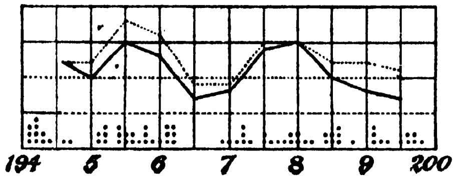
	Ne.	A.	Kr.	X.	RaEm.
ν_1	49·24	179·50	786·45	1777·90	5370·7
ν_2	19·71	75·60	309·20 341·16	815·05	2649·53
Δ_1	666	2519	10969	24893	72820
Δ_2	267·708	1057·057	4242·18 4678·80	10998·14	32166·44
δ	14·47013	57·9209	249·536	611·0100	1787·024
Δ_1/δ	46	40¾	44	40¾	40¾
Δ_2/δ	18½	18¼	17 18¾	18	18
e	191·5	719·71	3183·34	7314·1	23678·4
u	106·78 ?	439·47	1884·03	4133·18	11191·8
v	106·86 ?	442·67	1942·44	4428·00	13680·0
S(∞)	52112 52403	51731·05	51651·29	51025·29	50403·00

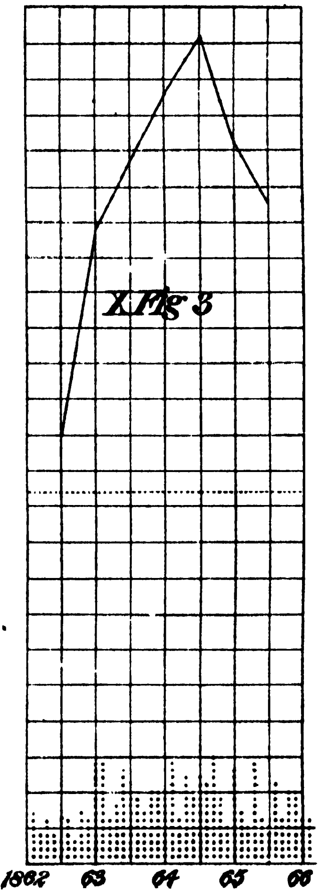
Kr Fig 1

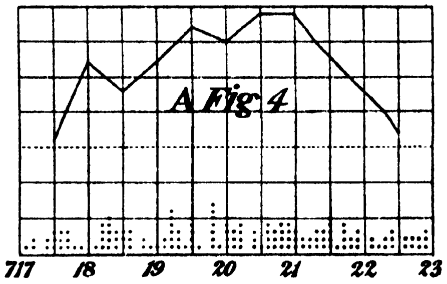

u e v

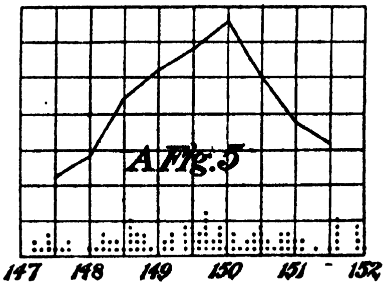

X Fig 2

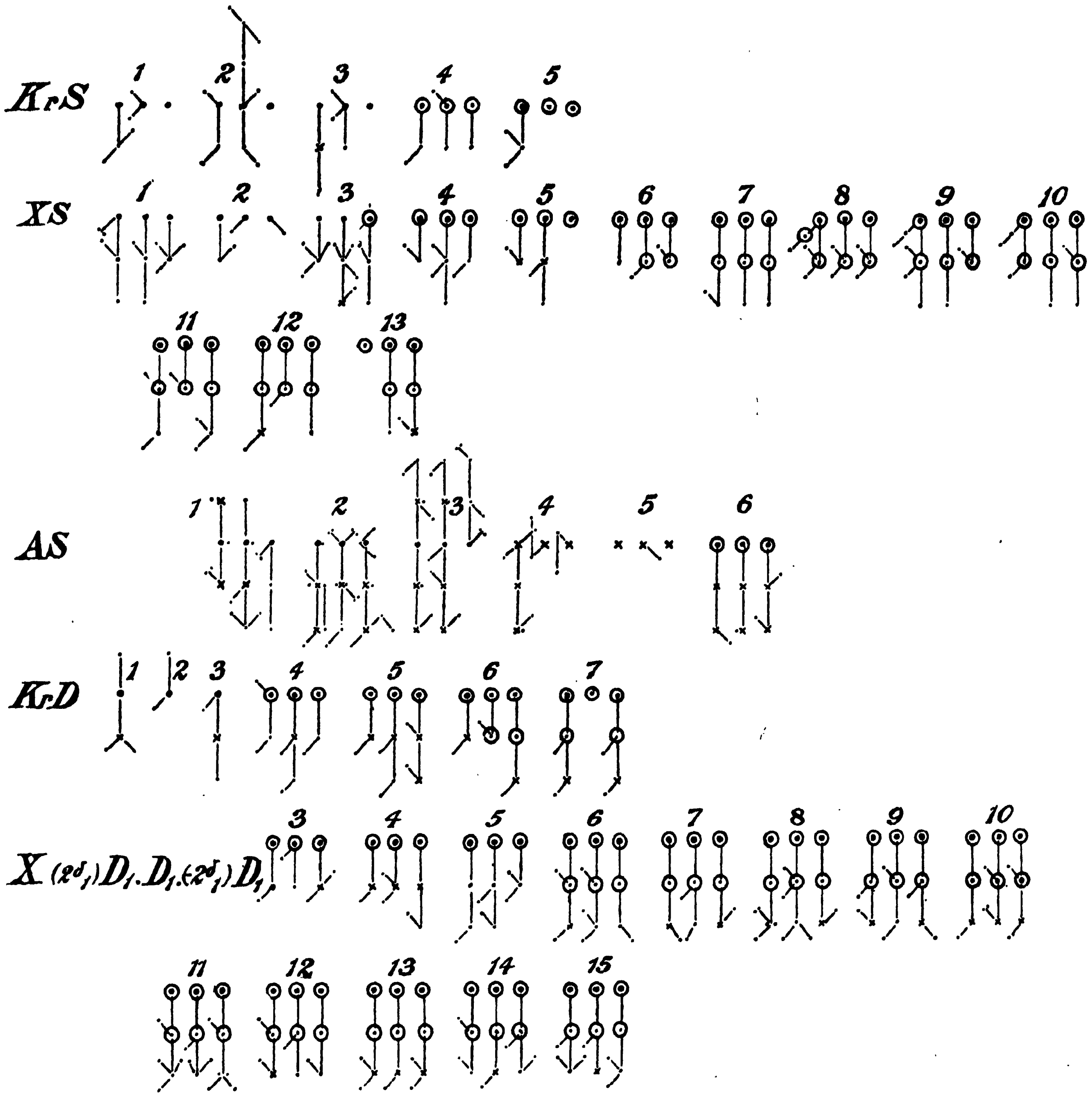
Ne Fig 6




X Fig 3


A Fig 4


A Fig 5



EXPLANATION OF DIAGRAM.

(1) A dot represents a wave-number; a large dot the wave-numbers being sounded for.

(2) A dot to the side of a line denotes a displaced line.

(3) A circle round a dot denotes that it is in an unobserved region.

(4) A × denotes that the wave-number has not been seen.

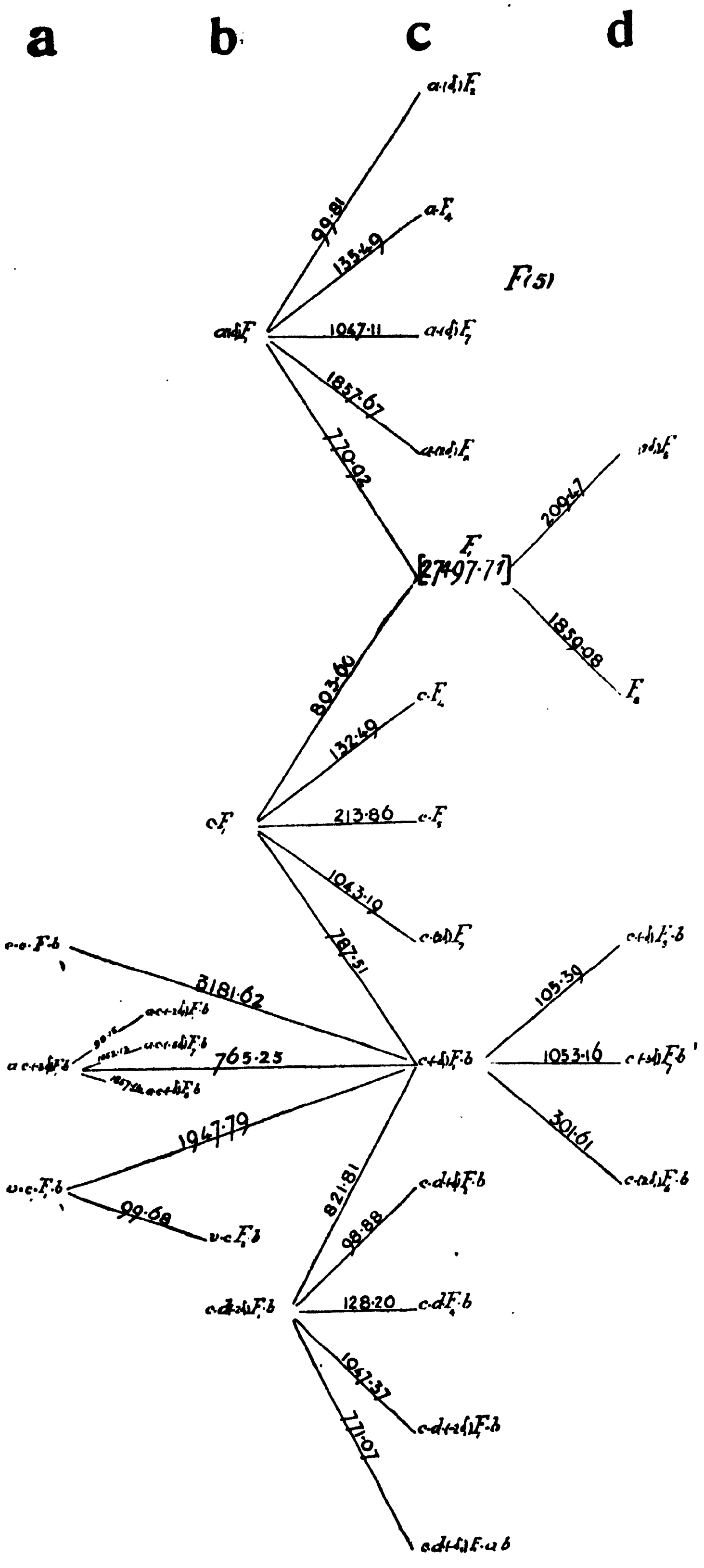
(5) The e links are represented by vertical lines, the u by lines at 45 degrees above the horizontal, the v by lines at 45 degrees below the horizontal.

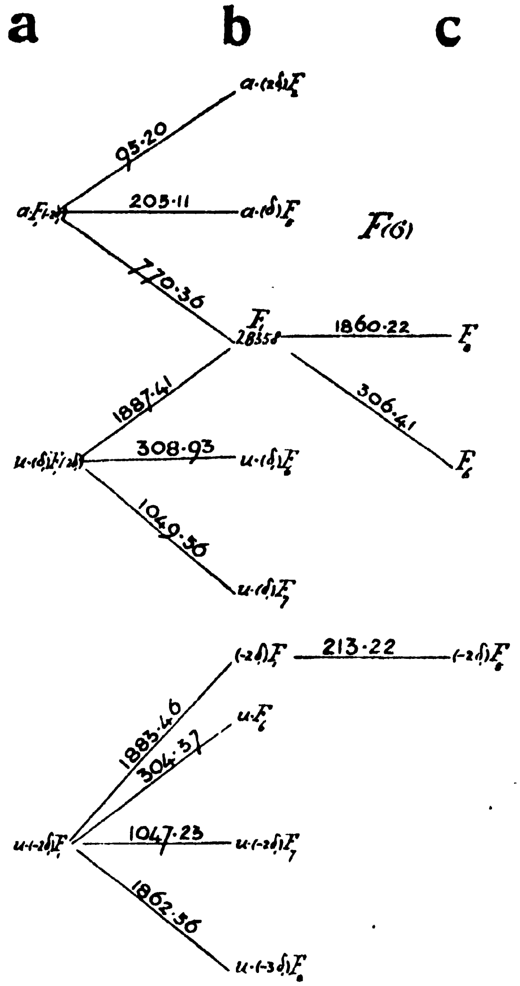
(6) The e link is to be subtracted when drawn down and added when drawn up. The u, v links are to be subtracted when drawn to the left and added when drawn to the right.

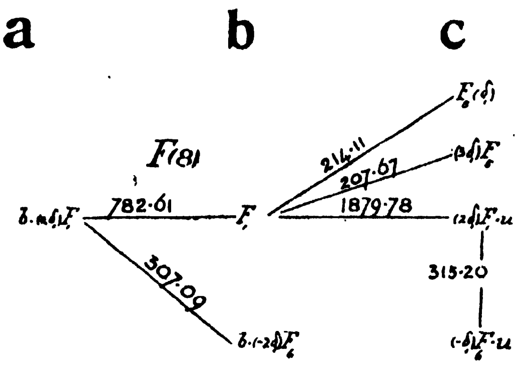
E.g., XS_2(3) is seen; also the lines linked to it by $-e$, $-e-u$, $-e+u$, $-e+v$; that linked by $-2e$ is not seen, but by $-2e+u$ is.

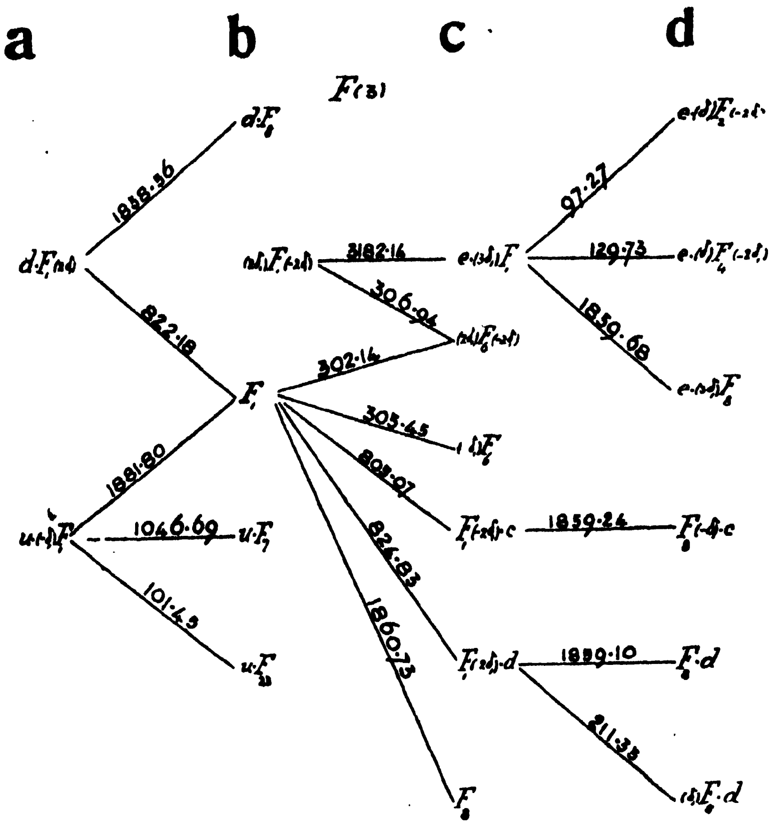
XS_3(3) is out of the observed region, but the lines linked to it by $-e$, $-e-u$, $-2e$ are in the observed region and have been seen.

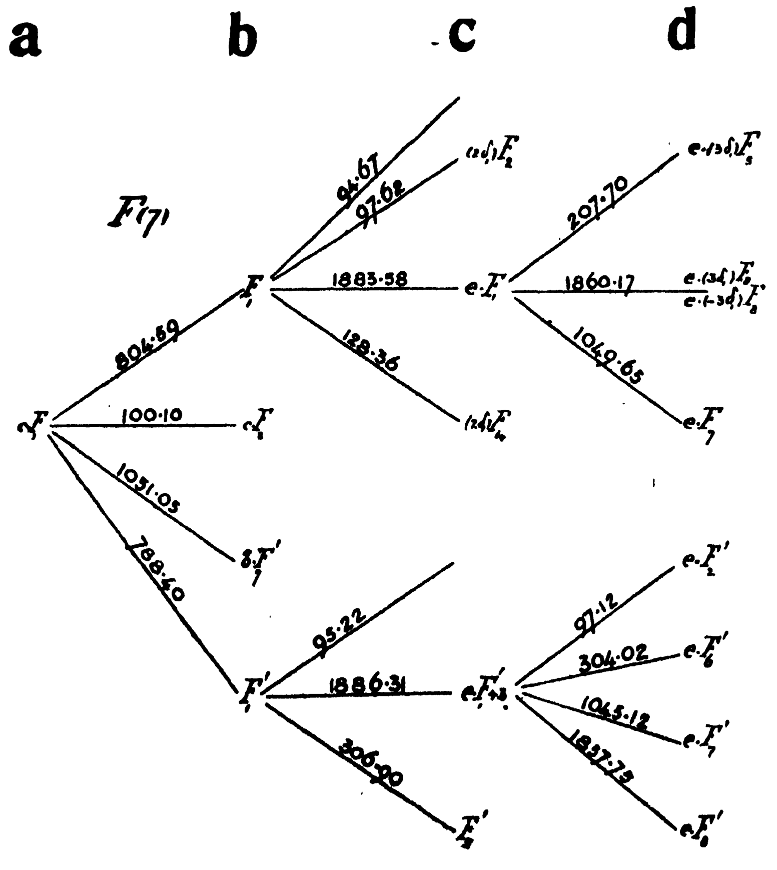
AS contains several examples of displaced lines.

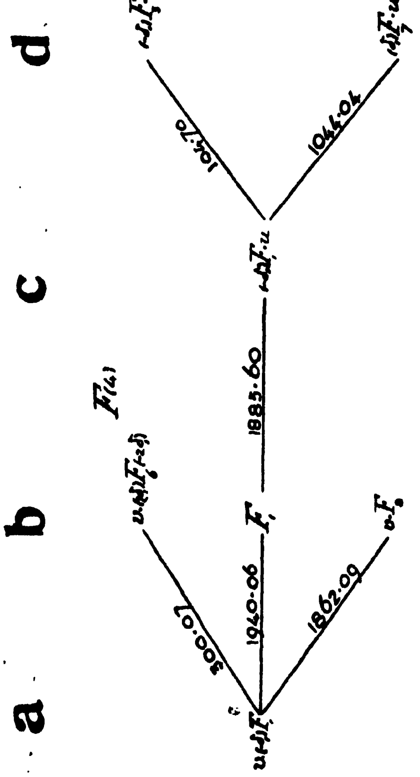

a
b
c
d
F(5)
99·81
135·49
1047·11
1837·67
770·92
F
[2797·71]
209·47
1859·08
803·60
132·49
213·86
1043·10
787·51
3181·62
765·25
1947·79
99·68
105·39
1053·16
301·61
821·81
98·88
128·20
1047·37
771·07


a
b
c
F(6)
95·20
205·11
770·36
F
28358
1860·22
306·41
1887·41
308·93
1049·56
213·22
1883·46
304·37
1047·23
1862·56


a
b
c
F(8)
782·61
307·09
214·11
207·67
1879·78
315·20


a
b
c
d
F(3)
d.F_8
1858·56
d.F_1(2δ)
822·18
F_1
1881·80
u(-δ)F_1
1046·69
u.F_7
101·45
u.F_12
(2δ_1)F_1(2δ)
3182·16
e(-3δ_1)F_1
306·04
302·14
(2δ_1)F_6(2δ)
97·27
e(δ)F_2(-2δ)
129·73
e(δ)F_4(-2δ)
1859·68
e(-2δ)F_8
305·45
(δ)F_6
808·07
F_1(2δ)c
1859·24
F_8(-4)c
824·83
F_1(2δ)d
1859·10
F_8 d
1860·73
211·38
(δ)F_8 d
F_8


a
b
c
d
F(7)
F_1
94·67
97·62
(2δ_1)F_2
1883·58
e.F_1
128·36
(2δ)F_4
207·70
e(-3δ_1)F_5
1860·17
e(-3δ_1)F_8
e(-3δ_1)F_8'
1049·65
e.F_7
804·59
F
100·10
F_8
1051·05
δ.F_7'
788·40
F_1'
95·22
1886·31
e.F_1+δ'
306·90
F_2'
97·12
e.F_2'
304·02
e.F_6'
1045·12
e.F_7'
1837·73
e.F_8'


a
b
c
d
F(4)
u(-δ)F_1
300·07
u(2δ)F_6(2δ)
1940·06
F_1
1862·09
u.F_8
1885·60
(-δ)F_1.u
104·70
(-δ)F_5.u
1044·04
(δ)F_7.u

INDEX

TO THE

PHILOSOPHICAL TRANSACTIONS.

SERIES A, VOL. 220.

B.

BAKERIAN LECTURE (PARSONS), 67.

BROMWICH (T. J. I'A.) The Scattering of Plane Electric Waves by Spheres, 175.

BURNHAM (T. H.). See DUFFIELD (and others).

D.

DAVIDSON (C). See DYSON (and others).

DAVIS (A. H.). See DUFFIELD (and others).

Diamond, artificial production of (PARSONS), 67.

DUFFIELD (W. G.) (and others). The Pressure upon the Poles of the Electric Arc, 109.

DYSON (Sir F.) (and others). A Determination of the Deflection of Light by the Sun's Gravitational Field, from Observations made at the Total Eclipse of May 29, 1919, 291.

E.

Eclipse of May 29, 1919 (DYSON and others), 291.

EDDINGTON (A. S.). See DYSON and others.

Electric arc, pressure upon poles of (DUFFIELD), 109.

Electric waves, scattering of plane, by spheres (BROMWICH), 175.

Electrodynamics, fundamental formulations of (LIVENS), 207.

Electromagnetic integrals (GREENHILL), 35.

G.

GREENHILL (Sir G.). Electromagnetic Integrals, 35.

H.

Helium, intensity relations in the spectrum of (MERTON and NICHOLSON), 137.

HICKS (W. M.). A Critical Study of Spectral Series. Part V.—The Spectra of the Monatomic Gases, 335.

I.

Integrals, electromagnetic (GREENHILL), 35.

[Published June 8, 1920.

L.

Light, determination of deflection of, by sun's gravitational field (Dyson and others), 291.
Livens (G. H.). On the Fundamental Formulations of Electrodynamics, 207.

M.

Magnetic susceptibility and molecular constitution (Oxley), 247.
Merton (T. R.) and Nicholson (J. W.). On Intensity Relations in the Spectrum of Helium, 137.
Molecular constitution and magnetic susceptibility (Oxley), 247.

N.

Nicholson (J. W.). See Merton and Nicholson.

O.

Oxley (A. E.). The Influence of Molecular Constitution and Temperature on Magnetic Susceptibility. Part IV.—Further Applications of the Molecular Field, 247.

P.

Parsons (Sir C. A.). Experiments on the Artificial Production of Diamond, 67.

S.

Spectral series, a critical study of—V. Monatomic gases (Hicks), 335.
Spectrum of helium, intensity relations in (Merton and Nicholson), 137.

T.

Taylor (G. I.). Tidal Friction in the Irish Sea, 1.
Tidal friction in the Irish Sea (Taylor), 1.

W.

Waves, scattering of plane electric, by spheres (Bromwich), 175.

HARRISON AND SONS, PRINTERS IN ORDINARY TO HIS MAJESTY, ST. MARTIN'S LANE, W.C. 2.

PHILOSOPHICAL TRANSACTIONS.

I. *Some Measurements of Atmospheric Turbulence.*

By Lewis F. Richardson.

Communicated by Sir Napier Shaw, *F.R.S.*

Received October 16, 1919,—Read February 26, 1920.

[Plate 1.]

Contents.

		Page
I.	Notation	1
II.	Shearing stress deduced from pilot balloon observations	3
III.	Eddy-viscosity from the same. It varies with height and with direction (anisotropic)	5
IV.	Eddy-diffusivity from the dispersal of smoke or of parachutes	5
V.	General theory of eddy-diffusivity deduced from scattering	8
VI.	Osborne Reynolds' eddy-stresses. A measure of gustiness	10
VII.	Summary of theory of scattering of particles of air	15
VIII.	Numerical values from scattering of particles	19
IX.	Cumulus eddies in calm weather	26
X.	Summary	28

I. Notation.

The following notation is used throughout. The co-ordinate axes are a right-handed rectangular system x, y, h, in which $0h$ is directed vertically upwards, and $0x$ lies in any azimuth which happens to be convenient. Elements of distance to east and to north are denoted by de, dn, so that they are special cases of dx, dy. The atmospheric density is ρ, the pressure is p, acceleration of gravity is g, latitude ϕ is reckoned negative in the southern hemisphere, and ω is the angular velocity of the earth. Velocities are denoted by v with a suffix to indicate the direction *towards* which they blow. Momenta per unit volume are denoted by m_X, m_Y, m_H. The eddy-diffusivity is denoted by a capital K as in G. I. Taylor's recent papers. Another, and in the author's opinion a better, measure of turbulence is ξ discussed in a previous paper.* The relation K to ξ is given by

$$\frac{\partial}{\partial p}\left(\xi \frac{\partial \chi}{\partial p}\right) = \frac{\partial \chi}{\partial t} = \mathrm{K}\frac{\partial^2 \chi}{\partial h^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where χ is either potential temperature, or else mass of water or smoke per mass of atmosphere. If ρ and ξ were independent of height, then from (1) we should have

$$\xi = g^2\rho^2\mathrm{K}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

---

* L. F. Richardson, 'Roy. Soc. Proc.,' A, vol. 96 (1919), pp. 9 to 13.

It is suggested that ξ might be named "the turbulivity." Its dimensions are: $(\text{mass})^2 \times (\text{length})^{-2} \times (\text{time})^{-5}$.

The advantage of using ξ instead of K is that the former enables one to allow for variations of density and of turbulence in a simple and natural manner. The disadvantage of ξ is that it has no name derivable from indoor physics. It is suited to the free atmosphere. We might compromise by using in place of K or ξ the "eddy-conductivity," c, defined by the equation

$$\frac{\partial(\rho\chi)}{\partial t} = \frac{\partial}{\partial h}\left(c\frac{\partial\chi}{\partial h}\right) \text{ or, approximately, } \frac{\partial\chi}{\partial t} = \frac{1}{\rho}\frac{\partial}{\partial h}\left(c\frac{\partial\chi}{\partial h}\right). \quad . \quad . \quad . \quad (3)$$

In so doing we gain an acceptable name "conductivity," but we lose by the explicit appearance of density in the equation. Either c or ξ allows variations of turbulence with height to be treated correctly, while K does not do so, as has been pointed out elsewhere by the author.* The dimensions of c are $(\text{mass}) \times (\text{length})^{-1} \times (\text{time})^{-1}$.

This c is of the same dimensions as the measure of turbulence discussed by W. SCHMIDT, of Vienna, under the name of "Austausch" in two important papers. ('Sitz. Akad. Wiss.,' Wien, 1917 and 1918.)

However much turbulence and density may vary with height

$$g^2\rho c = \xi. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

On the contrary if there are no variations with height,

$$c = \rho\mathrm{K}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

The six components of stress are denoted by $\widehat{xx}$, $\widehat{yy}$, $\widehat{hh}$, $\widehat{xy}$, $\widehat{yh}$, $\widehat{hx}$, as in the writings of K. PEARSON.

The convention adopted for the signs of eddy-stresses conforms to that of LOVE'S "Theory of Elasticity." Tractions are reckoned positive. That is to say, a direct stress such as $\widehat{xx}$ is positive if it be a tension, negative if a pressure; and a shearing stress such as $\widehat{xh}$ is positive when the air on that side of a level surface for which h is greater (*i.e.*, above), drags the air below in the sense of x increasing.

The *definition of eddy-viscosity* adopted in this paper is

$$\frac{\text{eddy shearing stress}}{\text{rate of mean shearing strain}}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\mathrm{A})$$

in agreement with the definition used by W. SCHMIDT (*loc. cit.*, 1917, p. 5).

The advantage of this definition is that it is simply based on the fundamental ideas of stress and strain, as well as being in harmony with the definition adopted in the theory of viscous liquids. (*cf.*, LAMB, 'Hydrodynamics,' IV. edn., § 326).

The question may arise as to whether the viscosity defined by (A) can ever become infinite by the vanishing of the denominator. The point is discussed by the author in the paper already cited (p. 13), and the conclusion is reached that such an occurrence would be highly improbable.

---

* *Loc. cit.*

The relation of ξ to the eddy-viscosity is most easily reached *via* terms $\partial m_x/\partial t$, $\partial m_y/\partial t$ in the dynamical equations. If pressure-gradient just balanced geostrophic wind we should have

$$\frac{\partial\,(\widehat{xh})}{\partial h} = \frac{\partial m_x}{\partial t} \qquad (6)$$

Now by the definition of viscosity given above

$$\widehat{xh} = \mu \, . \, \partial v_x/\partial h. \qquad (7)$$

Substituting (7) in (6) and inserting $m_x = \rho v_x$ there results

$$\frac{\partial}{\rho\,\partial h}\left\{\mu\,\frac{\partial v_x}{\partial h}\right\} = \frac{\partial v_x}{\partial t}. \qquad (8)$$

It is seen that this equation becomes identical with (1) if

$$\chi = v_x \quad \text{and} \quad \xi = g^2\rho\mu, \qquad (9)$$

of which the latter is the required relation.

On comparing equations (8) and (3), it is seen that eddy-viscosity, μ, and eddy-conductivity, c, are of the same dimensions, and appear in their respective differential equations in the same way. Indeed, TAYLOR has suggested that they are equal.* This likeness would be a good argument for recording observations in terms of these two quantities instead of in terms of diffusivity K or turbulivity ξ.

II. SHEARING STRESS FROM PILOT BALLOON OBSERVATIONS.

(*Condensed and revised January 22, 1920.*) EKMAN† in a remarkable paper pointed out that the total momentum of water produced by a tangential stress on the surface of the sea, in the steady state, is directed at right angles to the tangential stress, and its amount is quite independent of the value of the viscosity or of the variation of viscosity with depth. The same applies to the atmosphere. We may use this principle to find the shearing stress on the ground, provided we have a measure of what the momentum would be if the surface stress were zero.

I have taken the wind at a height of 1½ km. to 2½ km. as the standard of reference, because, by so doing, the term depending on curvature of path, and other small terms in the dynamical equations, are automatically allowed for to a first approximation. The stress at 2 km. is undoubtedly much less than that on the ground, and is neglected. It is best to select observations in which the momentum becomes nearly independent of height above 1·5 km. A table of results follows. They were computed with the help of Mrs. L. F. RICHARDSON. Dr. H. JEFFERYS says the selection will select abnormal lapse-rates and so abnormal viscosities.

---

* 'Phil. Trans.,' A, vol. 215, p. 22.

† "On the Influence of the Earth's Rotation on Ocean Currents," by V. W. EKMAN, 'Arkiv for Matem. Astr. och Fysik,' Stockholm, Bd. II., No. 11 (1905).

TABLE I.—Shearing Stresses at the Earth's Surface.

Here $\bar{m}$ is the mean momentum per cm.^{3} of the air between the ground and a height 2 km. above sea level.

Place.	Latitude.	$\bar{m}$. $\frac{\text{grm.}}{\text{cm.}^2\ \text{sec.}}$.	Resultant stress $\div \bar{m}^2$. $\frac{\text{cm.}^3}{\text{grm.}}$.	Stress on ground is veered* from wind near surface.	Stress dragging ground in the direction of $\bar{m}$. dynes cm.$^{-2}$.	Stress dragging ground perpendicularly to left of $\bar{m}$.	Reference and notes.
Eskdalemuir .	55°·3	0·72	6·7	−8°	+2·9	+1·9	Mean of 39 observations in 1913 and 1914. Light winds omitted.
Lindenberg .	52·2	1·24	1·0	26°	+1·5	+0·05	General mean. E. GOLD, Met. Office, 'Geophys. Mem.,' V., p. 143.
Upavon . .	51·3	1·70	1·0	assumed to be zero	+2·9	+0·8	Strong winds. DOBSON, 'Q. J. Met. Soc.,' April, 1914.
Batavia . . .	−6·2	0·26	1·3	−21°	+0·08	−0·03	May and June . . . . VAN BEMMELEN; Batavia, 'Verhandelingen,' vol. 1, 1911, pp. 13, 14.
		0·47	1·1	30°	+0·04	−0·24	July to September . .
		0·47	0·6	±180°	−0·12	−0·05	December to February .
The above were obtained by the method described in the preceding page. Below follows a result taken from G. I. TAYLOR'S paper, 1915.							
Upavon . . .	51·3	as above	0·83	assumed to be zero	—	—	The same observations as above.

---

* Veering here means turning in the sense of north to east, and this statement applies to both sides of the equator.

III. EDDY-VISCOSITY FROM PILOT BALLOON OBSERVATIONS.

(*Abridged, January 22, 1920.*) The method of the last section will give the difference of the shearing stresses on two surfaces of any horizontal slab of air. If we choose one of the surfaces so that $\partial v/\partial h = 0$ and consequently also the stress vanishes, we obtain the stress on the other surface. This has been done for some very smooth means for Lindenberg (E. GOLD, Met. Office, 'Geophys. Mem.,' V., p. 143). The results are set out in the following table :—

TABLE II.—Lindenberg.

The x axis is directed with the surface wind. The stress is that exerted by the upper on the lower layer.

Height above mean sea, kilometres.	Eddy-shearing-stresses.		Rates of mean shearing.		Eddy viscosities.	
	$\widehat{xh}$ dynes cm.$^{-2}$.	$\widehat{yh}$ dynes cm.$^{-2}$.	$\frac{\partial v_x}{\partial h}$ sec.$^{-1}$ $10^3 \times$	$\frac{\partial v_y}{\partial h}$ sec.$^{-1}$ $10^3 \times$	Parallel to wind. $\frac{\widehat{xh}}{\partial v_x/\partial h}$ dyne cm.$^{-2}$ sec.	Perpendicular to wind. $\frac{\widehat{yh}}{\partial v_y/\partial h}$ dyne cm.$^{-2}$ sec.
1·0	+0·05	−0·09	− 1·0	− 0·2	50	450
0·8	+0·01	−0·37	− 1·2	− 1·5	10	250
0·7	zero					
0·6	+0·02	−0·71	1	− 3·2	20	220
0·4	+0·09	−0·99	6·0	− 9·0	15	110
0·3	+0·39	−1·06	12·5	−12·2	31	87
0·2	+0·69	−0·90				
0·12 ground	+1·10	−0·64	21·5 ?	−17·5 ?	51	37

To obtain a quantity comparable with ξ we must multiply the eddy-viscosity by $g^2\rho$ which is approximately 1100 c.g.s. units.

The mean of the viscosities in the two directions increases with height as we might expect from other observations (*vide* Part VIII., below, also 'Roy. Soc. Proc.,' A, vol. 96 (1919), p. 18). But the most interesting thing about this table is the *marked lack of isotropy in viscosity. The air appears to be more viscous, for large motions, across the wind than parallel to it, except just near the ground.*

IV. EDDY-DIFFUSIVITY FROM SMOKE OR FLOATING BODIES.

Some direct measurements have been made by observing the gradually increasing scatter of smoke or other visible material carried along by the air. The changes in height of a large number of small portions of air are observed during a fixed interval of time. These changes are found to be distributed about their mean value approximately according to the ordinary "law of error." Their scatter in height is measured by the "standard deviation," computed by the familiar methods.* In order to find the diffusivity K the observations are compared with an appropriate integral of the approximate equation

$$\frac{\partial \chi}{\partial t} = \mathrm{K} \frac{\partial^2 \chi}{\partial h^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Such a one is

$$\chi = \frac{\mathrm{A}_1}{\sqrt{4t+\mathrm{A}_2}} e^{\frac{-(h-\mathrm{A}_3)^2}{\mathrm{K}(4t+\mathrm{A}_2)}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where A_1, A_2, A_3 are constants.

This integral represents a horizontal lamina in which the density χ is distributed about a mean height A_3 according to the law of error. The square of the standard deviation of the mass in the lamina can be shown to be

$$(4t+\mathrm{A}_2)\frac{\mathrm{K}}{2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

So if the scatter of the same set of particles be observed at the beginning and at the end of an interval T of time, it follows that

$$\mathrm{K} = \frac{1}{2\mathrm{T}} \text{ (increase during T of square of standard deviation).} \quad . \quad . \quad (4)$$

But the increase of the square of the standard deviation is equal to the square of the standard deviation of the change of height. Accordingly

$$\mathrm{K} = \frac{1}{2\mathrm{T}} \text{ (square of standard deviation of change of height during T).} \quad . \quad (5)$$

In this last transformation we have assumed that K is sensibly independent of height. This is permissible because the range of scatter can usually be made small. For the same reason the density of the air may be taken as independent of the height, so that we may obtain from K, the constant ξ, which we require when pressure is taken as independent variable in place of height, in accordance with (1) above. This procedure is not perfectly satisfactory but it is very convenient. It gives $\xi = g^2\rho^2\mathrm{K}$ and "eddy-conductivity" $= \rho\mathrm{K}$.

There is no need for the changes in height to be simultaneous for all the portions of air, and in practice it is much more convenient to let them be successive.

Varieties of particles.—I have observed the scattering of smoke from a smouldering wick, from burning weeds, from factory chimneys and from ship's funnels: also the scattering of portions of cloud near the horizon and of puffs of ammonium chloride from a special apparatus. Chimney smoke is not to be recommended, as it rises through the air. Clouds and steam may mislead one by evaporating. The cold NH_4CL smoke proved more satisfactory in these ways. Lycopodium dust might be better still, as isolated grains would fall at a definite rate relative to the air. But I have not succeeded in making the lumps break up into grains. The downy parachute which carries the seed of the dandelion, *Taraxacum officinale*, has been found to be convenient. When the seed is broken off, the parachute falls at a rate of 10 to 15 cm. sec.$^{-1}$ in still air. The standard deviation of this rate must be allowed for. The formulæ for correction are given below. A brown parachute, twice as large each way as that of taraxacum, grows near Benson. I am indebted to the Botanical Department of the British Museum for a search among their dried specimens for a large white parachute. A splendid one came from an African plant called strophanthus. The parachute is about 6 cm. in diameter and has a long stalk by which it can be held conveniently. When the seeds were broken off the parachutes fell, in still air, at an average rate of 20 cm. sec.$^{-1}$.

---

* *Vide* 'Computer's Handbook,' M.O. 223, Section V.

Other artificial clouds, which have been used with success, are paraffin-oil vapour from an extinguished blast lamp, and smoke of phosphorus pentoxide made by dropping calcium phosphide into dilute hydrochloric acid. In strong winds the smoke from a firework known as "Vesuvius"* is convenient.

If one could mark and follow individual molecules equation (5) would give the molecular diffusivity in still air, 0·2 cm.^{2} sec.$^{-1}$. Actually what we observe is the centre of a small puff of smoke, and this is not constantly the position of the same molecules, so that in still air we find $K = 0$. To be perfectly exact all observations of K by this method should be increased by 0·2 cm.^{2} sec.$^{-1}$, an entirely negligible correction. In any theory the diffusivity depends on the motions which the theory does not follow in detail. In laboratory experiments, in which the molecular motion only is ignored, K is taken as 0·2 cm.^{2} sec.$^{-1}$. In meteorological telegraphy variations of wind of less than 10 minutes' duration are ordinarily ignored, and there is an appropriate, much larger, value of the diffusivity. In a certain scheme for numerical prediction it is proposed to average the wind over periods of 6 hours, and the further variations thus omitted must be taken into account by further increase in K. It follows that the puffs of smoke should be so small as to allow the smallest eddies to be observed, and, for the last-named purpose, that the observations should be spread over a period of 6 hours. In obtaining the data in the following table I believe the former condition has been fulfilled, but the latter has not. When only the order of K or ξ is required, it is enough to assume that the standard deviation is $\frac{1}{5}$ of the distance between the extremes of height observed, when the number of observations is about 40.

Observations very near the earth's surface have peculiarities. It is obvious that $\partial\chi/\partial h = 0$ at an impermeable horizontal surface. This condition can be satisfied in the integral by taking the portion of the distribution which would be cut off by the surface, reflecting it in the surface, and adding it to the rest of the distribution.

---

* Made by Messrs. C. T. Brock & Co.

The standard deviation is not then sufficient to give K. It is necessary to make more elaborate computations with K. PEARSON'S "incomplete normal moment functions."

These difficulties are avoided in the symmetrical case when the source of smoke is exactly on the ground, and the smoke does not rise or fall by its temperature.

The results of these measurements are set out in the last column of Table IV. It is seen there that K increases from 5 near the surface of land up to 10,000 at the height of a factory chimney.

But before considering the results further a fuller mathematical investigation will now be made.

V. GENERAL THEORY OF EDDY-DIFFUSIVITY DEDUCED FROM SCATTERING.

The foregoing theory of the diffusion of a lamina assumes that the diffusivity is constant throughout the space, and that the density in the lamina does not vary except in the smooth regular manner indicated by the "law of error." But it is well known that the wind has an intricate structure. Thus if observations of the smoke puffs are to yield a measure of the diffusivity from the formula

$$\text{diffusivity} = \frac{\text{increase in square of standard deviation}}{\text{twice corresponding increase in time}},$$

then either the interval of time in the denominator must be long compared with the fluctuations of the wind in time, or the initial standard deviation must be large compared with the fluctuations of the wind in space, or both conditions must hold. The former condition is an inconvenient one in practice, because puffs are apt to fade before a sufficient time has passed. Dandelion parachutes, with the seeds removed, may be better than smoke for this purpose.

The following theory brings to light some of the assumptions involved in the measurement of diffusivity by smoke puffs. It was contrived specially in order to avoid "the distance through which an eddy moves before mixing with its surroundings," a quantity which occurs in TAYLOR'S theory, but which does not lend itself easily to measurement, except in the case of cumulus eddies. See Section IX. below.

The potential* temperature Θ does not change at a point moving with fluid, if radiation and precipitation can be neglected. Now let a portion of an eddy move from a height h_1 at time t_1 to a height h_2 at t_2. Then, regarding Θ as a function of h and t, we have

$$\Theta(h_1, t_1) = \Theta(h_2, t_2). \qquad \qquad (1)$$

---

* Potential temperature is the temperature which the air would acquire if compressed adiabatically to a standard pressure. If Θ is to be of service in dealing with cloudy air the standard pressure must be high enough to evaporate the cloud in all samples.

From each side of (1) subtract $\Theta(h_1, t_2)$ and divide through by t_2-t_1. Then

$$\frac{\Theta(h_1, t_1)-\Theta(h_1, t_2)}{t_2-t_1} = \frac{\Theta(h_2, t_2)-\Theta(h_1, t_2)}{t_2-t_1} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Now the left side of (2) is the finite difference ratio $\frac{\delta\Theta}{\delta t}$ at the height h_1, which is what we want to find in terms of the spacial distribution of Θ. Expand the right-hand side of (2) in powers of h_2-h_1 by the well-known theorem in the calculus. It follows that, if subscripts indicate the time and height

$$\left(\frac{\delta\Theta}{\delta t}\right)_{h_1, \frac{t_2+t_1}{2}} = \frac{1}{t_2-t_1}\left\{\left(\frac{\partial\Theta}{\partial h}\right)_{t_2 h_1}(h_2-h_1)+\left(\frac{\partial^2\Theta}{\partial h^2}\right)_{t_2 h_1}\frac{(h_2-h_1)^2}{2\,!}+\text{higher terms.}\right. \quad . \quad . \quad (3)$$

The difference ratio on the left of this equation is centred at the same height h_1 as the differential coefficients on the right of the same, but at a time $\frac{1}{2}(t_2-t_1)$ previous. This slight misfit in centering will not matter, because t_2-t_1 will be of the order of one minute or less, whereas we are next going to take the average of each term in (3) over a much longer time, say 6 hours. The subscripts may now be omitted as unnecessary. Let a bar over a symbol, or group of symbols, denote the mean value over this longer period. Let a dash denote the instantaneous deviation from this mean, so that, for instance, we have for every fluctuating quantity a formula such as

$$\frac{\partial\Theta}{\partial h} = \overline{\left(\frac{\partial\Theta}{\partial h}\right)}+\left(\frac{\partial\Theta}{\partial h}\right)' \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

Now the mean of any dashed quantity vanishes. . . . . (5)

Again the mean of the product of any dashed quantity into any barred quantity also vanishes. . . . . (6)

We shall further suppose that the mean of h_2-h_1 vanishes, ... (7)

that is to say that there is no mean vertical displacement.

Then, in the first term on the right of (3)

$$\begin{aligned}\overline{\frac{\partial\Theta}{\partial h}(h_2-h_1)} &= \overline{\left\{\overline{\left(\frac{\partial\Theta}{\partial h}\right)}+\left(\frac{\partial\Theta}{\partial h}\right)'\right\}\{(\overline{h_2-h_1})+(h_2-h_1)'\}} \\ &= \overline{\left(\frac{\partial\Theta}{\partial h}\right)'(h_2-h_1)'} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)\end{aligned}$$

because of (6) and (7).

Now $\overline{\left(\frac{\partial\Theta}{\partial h}\right)'(h_2-h_1)'}$, after being divided by t_2-t_1 and by the "standard deviations" of $\left(\frac{\partial\Theta}{\partial h}\right)'$ and of $(h_2-h_1)'$, becomes equal to the correlation between $\partial\Theta/\partial h$ and

$(h_2-h_1)/(t_2-t_1)$. So the first order term on the right of (3) vanishes, on taking the mean, if the variations of $\partial\Theta/\partial h$, in time at a fixed point, are not correlated with the variations in $(h_2-h_1)/(t_2-t_1)$, at the same point and time.

In cumulus cloud eddies, the variations of velocity are caused by variations of the potential temperature Θ, so that a correlation is almost certain to exist. On the contrary, when the eddies are due to dynamical instability, the correlation may be expected to vanish. In the latter case, it is the second order term of the right of (3) which becomes effective, so that

$$\frac{\overline{\delta\Theta}}{\overline{\delta t}} = \overline{\frac{\partial^2\Theta}{\partial h^2}\cdot\frac{(h_2-h_1)^2}{2(t_2-t_1)}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Now suppose further that either $\partial^2\Theta/\partial h^2$ has no variations at a fixed time and level, or else that its variations are not correlated with those of $(h_2-h_1)^2$. Then (9) simplifies to

$$\frac{\overline{\delta\Theta}}{\overline{\delta t}} = \overline{\frac{\partial^2\Theta}{\partial h^2}}\cdot\frac{\overline{\{(h_2-h_1)^2\}}}{2(t_2-t_1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Thus

$$\overline{(h_2-h_1)^2}/2(t_2-t_1) \text{ is the eddy-diffusivity K} \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

It is seen to be identical with that derived in Part IV. above, by considering the diffusion of a lamina, in which the density was distributed according to the law of error. It is a quantity easily measured.

Of course if t_2-t_1 were sufficiently small, say $\frac{1}{10}$ second, then it would be the first power of h_2-h_1, which would be proportional to t_2-t_1, instead of the square. This suggests that t_2-t_1 must be long compared with the fluctuations of the wind. On the other hand t_2-t_1 must be short compared with the period, of say 6 hours, over which the averages denoted by the bar are desired to be taken.

A similar argument can be applied to any other quantity which, like Θ, does not change following the motion of the fluid, provided it has space-rates independent of the time-variations of velocity. Thus the mass-of-water-per-unit-mass-of-atmosphere may replace Θ in (10) with similar restrictions.

When we consider diffusion in three dimensions there may be six coefficients of diffusivity corresponding to the six components of stress.

VI.—OSBORNE REYNOLDS' EDDY-STRESSES.

But we cannot, without further investigation, apply the preceding argument to the diffusion of horizontal velocity in a fixed azimuth.

Something might perhaps be deduced from the well-known theorem that, when ρ is constant,

$$\frac{D}{Dt}(\tfrac{1}{2}\rho v^2+\rho\psi+p) = \frac{\partial p}{\partial t} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

where ψ is the gravity potential.

The eddy-viscosity can however be measured rigorously by smoke-puff observations made in such a manner as to fit in with OSBORNE REYNOLDS' theory of eddy-stresses.* This theory is remarkably free from assumptions which might limit its generality. It is to be found in LAMB's 'Hydrodynamics,' IV. edn., Art. 369.

The equations of motion are three, such as

$$-\frac{\partial(\rho v_X)}{\partial t} = \frac{\partial p}{\partial x} + \frac{\partial}{\partial x}(\rho v_X v_X) + \frac{\partial}{\partial y}(\rho v_X v_Y) + \frac{\partial}{\partial h}(\rho v_X v_H)$$
$$+\rho\frac{\partial\psi}{\partial x} - 2\omega\sin\phi v_Y - c\left(\frac{1}{3}\frac{\partial \operatorname{div} v}{\partial x} + \nabla^2 v_X\right) \quad . \quad . \quad . \quad . \quad (13)$$

where c is the "molecular" or "ordinary" viscosity. Note that there is no need to assume ρ to be independent of position. REYNOLDS assumed this, but for a reason that does not concern us. It will be necessary however to assume that ρ', the variation of density at a fixed point, is so much smaller in comparison with $\bar\rho$ than is v' in comparison with $\bar v$, that we may put $\rho' = 0$. This being so, we find on taking the mean that (13) becomes

$$-\left\{\overline{\frac{\partial(\rho v_X)}{\partial t}}\right\} = \frac{\partial\bar p}{\partial x} + \rho\frac{\partial\psi}{\partial x} - 2\omega\sin\phi\bar v_Y - c\left(\frac{1}{3}\frac{\partial \operatorname{div}\bar v}{\partial x} + \nabla^2\bar v_X\right)$$
$$+\frac{\partial}{\partial x}(\rho\bar v_X \,.\, \bar v_X + \rho\overline{v'_X \,.\, v'_X}) + \frac{\partial}{\partial y}(\rho\bar v_X \,.\, \bar v_Y + \overline{\rho v'_X \,.\, v'_Y})$$
$$+\frac{\partial}{\partial h}(\rho\bar v_X \,.\, \bar v_H + \overline{\rho v'_X v'_H}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (14)$$

The left side of (14) is the difference between ρv_X at the beginning and at the end of the period through which the average is taken, divided by the period; and that is what we want. The right side of (14) is of exactly the same form in the mean quantities $\bar p$, $\bar v_X$, $\bar v_Y$, $\bar v_H$ as (13) was in the corresponding instantaneous quantities p, v_X, v_Y, v_H; except that there is added a force per unit volume in the x direction equal to minus

$$\frac{\partial}{\partial x}(\overline{\rho v'_X v'_X}) + \frac{\partial}{\partial y}(\overline{\rho v'_X v'_Y}) + \frac{\partial}{\partial h}(\overline{\rho v'_X v'_H}) . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

On working out the corresponding equations for the y and h components, it is seen that this additional force per unit volume is just that which would be given by the following systems of stresses

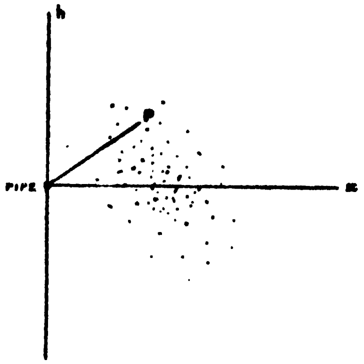
$$\left.\begin{array}{lll} \widehat{xx} = -\overline{\rho v'_X v'_X}\,; & \widehat{yy} = -\overline{\rho v'_Y v'_Y}\,; & \widehat{hh} = -\overline{\rho v'_H v'_H} \\ \widehat{xy} = -\overline{\rho v'_X v'_Y}\,; & \widehat{yh} = -\overline{\rho v'_Y v'_H}\,; & \widehat{hx} = -\overline{\rho v'_H v'_X} \end{array}\right\} \quad . \quad . \quad . \quad (16)$$

---

* Major G. I. TAYLOR tells me that he attempted to measure $\widehat{xh}$ with a balloon on an elastic tether in 1914.

when any symbol such as $\widehat{xy}$ is the force in the x-direction per unit area of a plane normal to the y axis. Tractions are reckoned positive, as usual. The above is taken from Osborne Reynolds' theory, adapted and slightly generalized to suit our needs for a rotating atmosphere, having density diminishing with height and a molecular viscosity which is not neglected.

One may form a clear mental picture of these eddy-stresses by imagining the scattering of smoke puffs. Let a puff emerge from a pipe at the origin of the co-ordinates. After a short interval of time τ, let the puff appear at the point P on the diagram, as seen by an observer at a distant point on the y-axis. Now let the observation be repeated for a large number of puffs in succession, the time τ being kept the same for each. We thus obtain a diagram, with a large number of points on it, showing the scattering of the puffs after τ. Then the eddy-stresses are simply related to the correlations and standard deviations of this scatter-diagram—under certain conditions. For let X, Y, H now mean the co-ordinates of any one of the dots on the diagram reckoned from the source of smoke.



Then the velocities of the corresponding puff were

$$v_X = \frac{X}{\tau}; \qquad v_H = \frac{H}{\tau}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (17)$$

provided the time τ was so short that, during it, the velocity may be regarded as uniform and in a straight line. (18)

Again, the velocity of the puff will be equal to that of the air which it has replaced provided the puff is at the same temperature as the air, and provided that the pipe points parallel to the Y-axis so that the impulse with which the puff leaves the pipe does not show in the projection on the plane XOH.

Let us suppose that a number of puffs, n in all, are observed. In order to correspond with the time-mean taken over 6 hours, which was used in deriving the eddy-stresses from the equations of motion, these n puffs should be spread uniformly over a similar interval.

From the scatter diagram we can compute first the mean velocities. For the mean velocities are

$$\bar{v}_X = \frac{\bar{X}}{\tau} = \frac{1}{\tau . n}\Sigma X; \qquad \bar{v}_Y = \bar{Y}/\tau = \frac{\Sigma Y}{\tau . n}, \quad . \quad . \quad . \quad . \qquad (19)$$

where Σ has the meaning :—take the sum of what follows it, for n puffs.

Then the normal eddy-stresses, $\widehat{xx}$ and $\widehat{hh}$, are found thus

$$\widehat{xx} = -\rho\overline{v'_X v'_X} = -\rho\,\overline{(v_X - \bar{v}_X)(v_X - \bar{v}_X)}$$

$$= -\frac{\rho}{\tau^2 n}\Sigma(X-\bar{X})(X-\bar{X}). \quad \dots \dots \quad (20)$$

So

$$\widehat{xx} = -\rho\sigma_X^2/\tau^2, \quad \dots \dots \quad (21)$$

where σ_X is the "standard deviation" of the dots in the x-direction. If σ_H is the corresponding quantity vertically

$$\widehat{hh} = -\rho\sigma_H^2\tau^{-2}. \quad \dots \dots \quad (22)$$

So the direct eddy-stress in the direction of the wind is intimately related to the gustiness shown by a tube-anemometer.

The shearing eddy-stress

$$\widehat{xh} = -\rho\overline{v'_X v'_H} = -\frac{\rho}{\tau^2}\frac{1}{n}\Sigma(X-\bar{X})(H-\bar{H}). \quad \dots \dots \quad (23)$$

So

$$\widehat{xh} = -\rho\tau^{-2} r_{XH}\sigma_X\sigma_H, \quad \dots \dots \quad (24)$$

where r_{XH} is the correlation between the co-ordinates X and H of the dots.

By projecting the puffs on the other two co-ordinate planes we should be able to measure similarly the remaining components of eddy-stress.

To find the eddy-viscosity we must compare the sheering eddy-stress $\widehat{xh}$ with $\left(\frac{\partial\bar{v}_X}{\partial h} + \frac{\partial\bar{v}_H}{\partial x}\right)$, which is the rate of shearing strain in the *mean* motion. Usually $\partial\bar{v}_H/\partial x$ is negligible, so that the rate of shearing is $\partial\bar{v}_X/\partial h$, a quantity which can easily be observed. At first sight one might think that $\partial\bar{v}_X/\partial h$ was simply related to the slope of the regression line in the scatter diagram; but on examination this proves not to be the case. The slope of the regression line is independent of τ, because (18) is satisfied for all permissible intervals of time.

It should be noted that no shearing stress such as $-\rho\overline{v'_X v'_H}$ can exceed, in absolute value, the geometric mean of the corresponding pair of direct stresses $-\rho\overline{v'_X v'_X}$, $-\rho\overline{v'_H v'_H}$ for the same reason that a correlation coefficient cannot exceed unity.

The probable errors of eddy-stresses, determined from the scattering of particles moving with the air, may be taken to be as follows

$$\text{Probable error of } \widehat{xx} = 0{\cdot}674\,.\,\widehat{xx}\sqrt{\frac{2}{N}}. \quad \dots \dots \quad (25)$$

$$\text{Probable error of } \widehat{xh} = 0{\cdot}674\sqrt{\frac{\widehat{xh}^2 + \widehat{xx}\,.\,\widehat{hh}}{N}}. \quad \dots \dots \quad (26)$$

These are based on the assumption that the number of particles is not less than say 20, and that the scatter is "normal," so that the logarithm of the density of particles would be a quadratic function of x, y, h if a very large number of particles were observed. Reference may be made to the fundamental papers on probable errors: FILON and PEARSON, 'Phil. Trans.,' A, vol. 191; W. F. SHEPPARD 'Phil Trans.,' A, vol. 192. Equation (26) follows from the probable error of the quantity called by K. PEARSON the "product moment coefficient *taken about the mean.*" Equation (25) follows from (26) on putting $x = h$, or may be deduced independently from the probable error of a standard deviation taken about the mean.

Corrections for the Motion of the Parachute Relative to the Air.

When observing eddy-stresses by the aid of the parachutes of plant seeds it is desirable to allow for the velocity of the parachute in still air. For large specimens of the parachute of *Taraxacum officinale*, after cutting off the seed, the velocity in still air was found to have a mean of 12 cm. sec.$^{-1}$ with a standard deviation of 2 cm. sec.$^{-1}$. It may be shown from the equation of motion that this limiting velocity is acquired in a negligibly short interval of time. Thus call the limiting velocity c downwards, the instantaneous velocity downwards u. Then if the friction is proportional to the velocity

$$g\,(\text{mass}) = c \times \mathrm{F}, \qquad \qquad (27)$$

where F is a constant. But, when accelerating,

$$(\text{mass})\,.\,\frac{du}{dt} = g\,(\text{mass}) - u\mathrm{F}. \qquad \qquad (28)$$

Eliminate the mass between these two equations and there results

$$\frac{c}{g}\,\frac{d\,(u-c)}{dt} = -\,(u-c). \qquad \qquad (29)$$

So that the discrepancy between the actual velocity u and the terminal velocity c sinks to e^{-1} of itself in a time equal to c/g, which for the taraxacum parachute having $c = 12$ cm. sec.$^{-1}$ would be only a hundredth of a second.

Less negligible is the variation of the velocity c from one parachute to another. What we actually observe is not v_{H} the upward velocity of the air, but $v_{\mathrm{H}}-c$. Now write $\bar{c}$ for the mean velocity of the parachute in still air, and c' for the deviation from the mean. Then in finding the direct stress $\widehat{hh}$ we must perforce work out first the "raw" moment $\{\overline{(v_{\mathrm{H}}-c)^2}\} = \overline{[\{(\bar{v}_{\mathrm{H}}-\bar{c})+(v'_{\mathrm{H}}-c')\}^2]}$. On expanding and remembering that a bar put over the product of a dashed and a barred symbol, causes the result to vanish, and also that $\overline{v'_{\mathrm{H}}-c'} = 0$, it is found that

$$\{\overline{(v_{\mathrm{H}}-c)^2}\} = (\bar{v}_{\mathrm{H}}-\bar{c})^2 + \overline{v'_{\mathrm{H}}v'_{\mathrm{H}}} + \overline{c'c'}$$

it follows that the corrected stress is given by

$$\widehat{hh} = -\rho\overline{v'_H v'_H} = -\rho\left\{\overline{(v_H-c)(v_H-c)}-(\bar{v}_H-\bar{c})^2-\overline{c'c'}\right\} \quad . \quad . \quad . \quad . \quad . \quad (30)$$

which is the actual formula used in working up the observations. Here $\overline{c'c'}$ is the square of the standard deviation of the velocity of the parachutes in still air.

Next we require the product moments in order to find the shearing stresses $\widehat{xh}$ and $\widehat{yh}$. The raw moment $\overline{v_X(v_H-c)}$ is found on expanding to be equal to $\bar{v}_X(\overline{v_H-c})+\overline{v'_X v'_H}$.

So that the corrected value for the stress is

$$\widehat{xh} = -\rho\overline{v'_H v'_X} = -\rho\left\{\overline{v_X(v_H-c)}-\bar{v}_X(\overline{v_H-c})\right\}, \quad . \quad . \quad . \quad . \quad . \quad . \quad (31)$$

and as $\overline{c'c'}$ does not appear, the scatter of the velocities of the parachutes in still air does not make a correction necessary for the shearing stresses.

VII. Summary of Theory of Scattering of Particles of Air.

The conclusion we have reached is the following. For any sort of eddy, whether due to "dynamical instability," or to the rising of heated air in cumuli, the eddy-stresses are best measured by equations (22), (24) and the like, because the theory from which they are derived is very general; and the eddy-viscosity is best measured as the ratio of the shearing eddy-stress to the rate of mean shearing strain. It is conceivable that $\widehat{xh}$ found from (24) might turn out to be zero. In that case it would be necessary to investigate effects of higher order. This might possibly be done by developing, for the quantity $(\frac{1}{2}\rho v^2+\rho\psi+p)$ in equation (12) an analysis similar to that of (1) to (11) for potential temperature. The diffusivity for potential temperature, on the other hand, should be measured differently according as the eddies are produced by variations of potential temperature or not. Thus for cumulus eddies we should take the mean of (3), retain the linear term on its right-hand side and neglect the quadratic one. Then t_2-t_1 must be small, so that

$$\overline{\frac{\delta\Theta}{\delta t}} = \overline{\frac{\partial\Theta}{\partial h}\cdot\frac{(h_2-h_1)}{t_2-t_1}} = \overline{\left(\frac{\partial\Theta}{\partial h}\right)' v'_H}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (31\text{A})$$

The diffusivity is measured as the right side of this equation divided by $\overline{\partial^2\Theta/\partial h^2}$. Thus

$$K = \frac{\overline{\left(\frac{\partial\Theta}{\partial h}\right)' v'_H}}{\overline{\partial^2\Theta/\partial h^2}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (32)$$

But for eddies due to dynamical instability, neglect the linear term in (3) and measure the diffusivity as

$$K = \frac{\overline{(h_2-h_1)^2}}{2(t_2-t_1)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (33)$$

where (t_2-t_1) must not be too small. It will be interesting to see whether eddy-diffusivity is found to be equal to eddy-viscosity divided by density.

Let us now see how these two modes of disposal of smoke fit together—the one, at short intervals of time, governed by the eddy-stresses, the other at long intervals governed by the diffusivity. Let us draw the trail of smoke as it would appear to a distant observer looking, say, horizontally. The scatter may be represented by drawing two lines through the points where the smoke-density has its standard deviation in height above or below its mean. Let T be the time since the smoke emerged from its source. Suppose that the stress $\widehat{hh}$, the diffusivity K and the mean velocity $\bar{v}_x$ are all constant along the path of the smoke. The time T taken to travel a horizontal distance x measured down the trail from the source is $x/\bar{v}_x$. Near the origin $T = \tau$ in equation (22), and so the standard deviation in height is

$$\pm \frac{x}{\bar{v}_x} \sqrt{-\widehat{hh}/\rho} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (34)$$

representing a pair of straight lines intersecting at the source. Further down the trail $T = t_2 - t_1$ in equation (11), and the standard deviation in height is

$$\pm \sqrt{x} \cdot \sqrt{\frac{2K}{\bar{v}_x}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (35)$$

representing a parabola with horizontal axis and its apex at the source. Thus according to the theory, the smoke may be said to be contained within a paraboloid which has had its blunt end sharpened into a cone. The preceding theory gives us no clue as to the manner of transition from the cone to the paraboloid.

When we can observe a sufficient length of *the path traced out in space by a single small portion of air* the eddy-stresses and the eddy-diffusivity may be deduced from the irregularities in the motion. With this object I have observed the motion of anti-aircraft shell bursts, and of portions of cloud, by means of an Abney level or a pocket sextant. With better instruments this method might yield a good deal of information about eddies at heights such as 2 to 5 km. One principal difficulty is that the shell-burst fades away after about 5 minutes, before a sufficient length of path has been observed to give the diffusivity.

If the path is sufficiently high and long, the hills, trees and houses on the earth may be regarded as blending into a "roughness." Suppose this roughness to be uniform. Then if we had been causing smoke to issue in puffs from a fixed pipe, we should presumably have obtained the same scatter diagram for the puffs, within the limits of probable error, at whatever point of the path we had placed the pipe, or at whatever time we had begun to observe, within limits. If this is so, we may form the scatter diagram by taking its origin at every point in succession of the trajectory of the single particle. For instance, Captain CAVE in his book on "The Structure of the Atmosphere," gives several diagrams of the irregularities of height of a balloon observed by two theodolites, the uniform vertical motion of the balloon relative to the air having been eliminated. From his figure 30, of an ascent on February 19, 1909, the following has been deduced, by taking the origin of scattering at every available minute mark, and working out the standard deviation, in height about the mean height, after a time α.

TABLE III.

α, secs.	1×60	3×60	5×60	7×60	9×60
$\frac{\sigma_H^2}{2\alpha} = 10^4 \times$ cm.^{2} sec.$^{-1}$	0·93	0·82	0·55	0·55	0·33
$\rho\frac{\sigma_H^2}{\alpha^2}$ dyne cm.$^{-2}$	0·345	0·100	0·041	0·029	0·014
Number of points	14	12	10	8	6

Now $\sigma_H^2/2\alpha$ would be the diffusivity if α were a "long" time; and the criterion of sufficiency in length is that $\sigma_H^2/2\alpha$ should not vary with α. The small number of points makes the probable errors large, so that the decrease of $\sigma_H^2/2\alpha$ between $\alpha = 7$ mins. and $\alpha = 9$ mins. is not significant. It looks as though the diffusivity K were here of the order of $0{\cdot}5 \times 10^4$ cm.^{2} sec.$^{-1}$.

Again $-\rho\sigma_H^2/\alpha^2$ would be the eddy-stress, $\widehat{hh}$, if α were so small that further decrease made no further change in the quantity. This stage is not reached at $\alpha = 1$ minute. All that we can say is that the eddy-stress is probably greater than 0·345 dynes cm.$^{-2}$.

The mean height of this observation is 1 km. above ground, and the mean velocity 1270 cm. sec.$^{-1}$.

The photograph of a smoke trail in fig. 2 suggests that the eddying is partly random but also partly sinusoidal. Let us therefore see what would happen *if the path of the particle were an exact sine curve* without any random variations. Let the height h of the particle be given by $h - \mathrm{B} = \mathrm{A} \sin qt$ where A, B and q are constants. The increase in height in a time α would be

$$\mathrm{A}\left[\sin\left\{q\left(t+\frac{\alpha}{2}\right)\right\} - \sin\left\{q\left(t-\frac{\alpha}{2}\right)\right\}\right] = \mathrm{A} \,.\, 2\cos qt \,.\, \sin\left(q\frac{\alpha}{2}\right)$$

by trigonometry. So that the standard deviation σ_H after a time α would be given by

$$\sigma_H^2 = \frac{4\mathrm{A}^2\left(\sin\frac{q\alpha}{2}\right)^2}{\mathrm{L}} \int_{t=0}^{t=\mathrm{L}} (\cos qt)^2\, dt$$

where L is a very long time. The integral is equal to $\frac{1}{2}$L plus a negligible oscillatory part. Consequently $\sigma_H^2 = 2\mathrm{A}^2\left(\sin\frac{q\alpha}{2}\right)^2$. It follows that the stress $\widehat{hh}$, which is the limit of $-\rho\sigma_H^2/\alpha^2$ when α is small, comes to $-\rho q^2\mathrm{A}^2/2$; whereas the diffusivity K, which is the limit of $\sigma_H^2/2\alpha$ when α is long, comes to zero.

Comparison with TAYLOR'S *expression for the Diffusivity and with* W. SCHMIDT'S *"Austausch."*

In TAYLOR'S remarkable investigation ('Phil. Trans.,' A, vol. 215) from which the present research took its stimulus, the diffusivity K is given, in the present notation, as the mean value of $v_H(h-h_0)$ over a large horizontal plane; and it is stated that $h-h_0$ is the height through which an eddy moves from the layer at which it was at the same temperature as its surroundings, to the layer with which it mixes. This definition of $h-h_0$ is puzzling, for it seems impossible to reconcile the supposed starting and stopping of the air, with the ceaseless motion which we observe in nature, except in the case of cumulus eddies. Happily we are now in a position to clear away the mystery. For it has been shown independently in the present paper that the diffusivity is given by

$$K = \overline{(h_2-h_1)^2/2(t_2-t_1)}$$

where (t_2-t_1) is a time long compared to the fluctuations in the wind, and where the bar implies an average taken over a still much longer time. As (t_2-t_1) is the same for all the quantities which are averaged, we may remove it from under the bar, writing

$$K(t_2-t_1) = \tfrac{1}{2}\overline{(h_2-h_1)^2}.$$

Differentiate this equation with respect to t_2,

$$K = \overline{(h_2-h_1)\frac{d}{dt_2}(h_2-h_1)} = \overline{(h_2-h_1)\,.\,v_H} \textit{ at } t_2,$$

thus K is expressed as the mean of the product of the rise in height during a long time into the vertical velocity *at the end* of that time. It may also be taken at the beginning. Comparing with TAYLOR'S form quoted above we see a strong resemblance, and we are led to suppose that TAYLOR'S theory makes two unnecessary and unnatural restrictions: (1) that the portion of air should start at the same temperature as its surroundings; (2) that the portion of air should finally mix with its surroundings. But that if these restrictions be removed, then another becomes necessary, namely that (t_2-t_1) should be sufficiently long (several minutes). Whether the average be taken over a large horizontal plane, or over a very long time (6 hours), appears to be a matter of indifference.

The extent to which TAYLOR assumes viscosity to be independent of height in his *general* theory ('Phil. Trans.,' A, vol. 215, pp. 11 to 13) is this: he neglects the terms due to the initial eddying in his equation (6). That is a doubtful proceeding, unless the initial eddying is zero: but zero is independent of height.

The "Austausch" of W. SCHMIDT is defined by him (in 'Sitz. Akad. Wiss.,' Wien (1917), pp. 4 to 5) as

$$\frac{\Sigma\,(\text{element of mass crossing horizontal plane})\times(\text{vertical displacement of element})}{(\text{whole area})\times(\text{time of motion})}.$$

If we replace the summation by an integration over a large area S in the plane of xy, the element of mass crossing per unit time is $dx\,dy\,\rho v_{\mathrm{H}}$, so that the "Austausch" becomes $\frac{1}{\mathrm{S}}\iint \rho v_{\mathrm{H}}(h-h_0)\,dx\,dy$ which is ρ times TAYLOR's diffusivity as defined in 'Phil. Trans.,' A, vol. 215, p. 3. Thus we must suppose that W. SCHMIDT's definition of "Austausch" requires amplification concerning the interval of time and concerning the position in it of the velocity, just as TAYLOR's definition of K does.

VIII. NUMERICAL VALUES DERIVED FROM THE SCATTERING OF PARTICLES.

Fig. 1 is a photograph* of the trail of paraffin vapour from an extinguished blast-lamp which projected the vapour in a direction at right angles to the wind. It shows a cone, with a blunt point due to the finite size of the source of smoke, passing smoothly into a form, which certainly diverges less rapidly than the initial cone, and which looks like a paraboloid. Opinion might differ slightly as to where to draw the lines corresponding to the standard deviation of smoke. In a "normal" distribution 0·68 of the whole number of particles lie between the two standard deviations. If the lines are placed as in the accompanying black and white drawing, then it follows, as the mean velocity of the smoke was 1·7 metres/sec., and the density of the air was $1{\cdot}21\times10^{-3}$ grm. cm.$^{-3}$, that

$$\text{stress}\ \widehat{hh} = -0{\cdot}73\ \text{dyne cm.}^{-2}\ \text{diffusivity} = \mathrm{K} = 240\ \text{cm.}^2\ \text{sec.}^{-1}$$

$$\text{turbulivity} = \xi = 340\ \text{grm.}^2\ \text{cm.}^{-2}\ \text{sec.}^{-5}.$$

This photograph was taken in the evening, when the day-wind was diminishing. The source was 190 cm. above ground. Obstructions to windward only subtended an angle of 2°·1 at the source of smoke. The exposure lasted 60 seconds.

Fig. 2 was taken five minutes later in the same place, with an exposure of 85 seconds.

The velocity of the smoke had decreased to 1·3 metres per sec. The measurements yield

$$\widehat{hh} = -1{\cdot}2\ \text{dynes cm.}^{-2};\qquad \mathrm{K} = 750\ \text{cm.}^2\,\text{sec.}^{-1};\qquad \xi = 1050\ \text{grm.}^2\,\text{cm.}^{-2}\,\text{sec.}^{-5}.$$

In this case the photograph shows a distinct neck between the cone and the paraboloid, at a distance from the source roughly 1·3 times its height above ground. This neck can also be recognized in some other photographs. Its presence signifies that the motion of the air was compounded of (i.) a random eddying, plus (ii.) a wave motion in which the particle of air executed a wave having a length, relative to a point fixed to the earth, roughly 2·6 times the height of the particle above ground.

---

[*] Taken at Benson.

TABLE IV.—Observed Values of

The co-ordinate axes are taken to point: $0x$ horizontally with the mean wind at
quantities are in C.G.S. units. Tractions are reckoned positive, so that $\widehat{xh}$ and $\widehat{yh}$

Entropy gradient vertically upwards.	Surface.	Methods.	Notes.	h cm. above surface.	$\bar{v}_x$ $\frac{\text{cm.}}{\text{sec.}}$	$\widehat{xx}$ dynes per square centimetre.
Positive sunrise . .	Tall grass . . .	Smoking wick . . .	I.	30 170 300,000	10 60 200	—
Small ?	Oat field . . .	NH_4Cl puffs . . . . NH_4Cl puffs . . . . Dandelion down . .	II.	40 100 50	20 40 —	— — —
Small ?	Corn field . . .	Smoke puffs	III.	120	100	− 0·1
Small ?	Moor	Smoking wick . . .	IV.	165	100	—
?	Moor, trees . .	Thistle down . . .	V.	200	145	− 2·4 ± 0·5
Positive ?	Flat field . . .	Paraffin vapour. . .	VI. Figs. 1 and 2	190	170 140	— —
1919, Mar. 4 Between England and the Isle of Wight.	Sea	Steamer's smoke . .		4,000	100 ?	—
1919, Apr. 7, 18h.	Sea	Steamer's smoke . .		10,000	200 ?	—
1919, July 4, 11h. overcast	Sea	Steamer's smoke . .		2,000	130	—
1919, July 5, 12h.	Sea	Steamer's smoke . .		1,400	140	—
	Ditcham . . .	Capt. CAVE's balloon .	p. 16	100,000	1270	—
	Flat fields . . .	Phosphorus pentoxide	VI. Fig. 3	340	120	—

Eddy-stresses and of *Eddy-diffusivity.*

the level of observation, 0y horizontally to the left and 0h vertically upwards. All are positive when the air above drags the air below in senses of x and y increasing.

$\widehat{yy}$ dynes per square centimetre.	$\widehat{hh}$ dynes per square centimetre.	$\widehat{xy}$ dynes per square centimetre.	$\widehat{yh}$ dynes per square centimetre.	$\widehat{hx}$ dynes per square centimetre.	$\frac{\partial \bar{v}_x}{\partial h}$ sec.$^{-1}$.	Conductivity $= \rho K$ grm. cm.$^{-1}$ sec.$^{-1}$.	Diffusivity. K cm.^{2} sec.$^{-1}$.	Turbulivity. $\xi = g^2\rho^2 K$ grm.^{2} cm.$^{-2}$ sec.$^{-5}$.
—	—	—	—	—	>0·5 — —	0·006 — —	5 — —	7 — —
— — —	-0·004 -0·006 —	— — —	— — ·	— — —	}0·3 { — —	— — 0·07	— — 60	— — 80
—	-0·2	—	—	<0·1	<0·06			
—	-0·04	—	—	—	—	0·03	24	34
- 2·9 ± 0·6	-0·6 ±0·12	[+0·45] ±0·39	-0·48 ±0·20	[-0·34] ±0·18				
— —	-0·7 -1·2	—	—	—	—	0·3 0·9	240 750	340 1,050
—	—	—	—	—	—	12	10,000	14,000
—	—	—	—	—	—	0·25	200	300
—	—	—	—	—	—	6·8	5,500	8,000
—	—	—	—	—	—	5·9	4,800	7,000
—	< -0·34	—	—	—	—	6	5,000	6,000
—	—	—	—	—	—	0·16	130	200

TABLE IV.—Observed Values of

Entropy gradient vertically upwards.	Surface.	Methods.	Notes.	h cm. above surface.	$\bar{v}_x$ $\frac{\text{cm.}}{\text{sec.}}$.	$\widehat{xx}$ dynes per square centimetre.
1919, Oct., 21d. 12h., BENSON	Obstructions up wind $\frac{1}{10}$ radian	Paraffin vapour. . .		160	250	—
1919, Oct., 29d. 15h., BENSON; raining	Obstructions up wind 0·045 radian	Smoke		190	330	—
Small ?	Moor, trees . .	Burning rubbish . .	VII.	900	300	—
Small ?	Fields, trees . .	Factory chimney . .	VIII.	3,000	600	—
Negative ?	Moor, low hills .	Cloud	IX.	150,000 2,000	1000 200	— —
Small ?	Wooded hills . .	Large fire	X.	25,000 37,000	— 1600	—
		Shell-puffs	XI.	300,000	500	—
Positive ?	Spurn head . .	Dines anemometer .	XII.	—	1200	-80
Positive ?	Open sea . . .	Steamer's smoke . .	XIII.	1,500	—	—
Negative ?	Open sea . . .	Steamer's smoke . .	XIV.	5,000	200	—
?	Open sea . . .	Alto-stratus	XV.	500,000	—	—

Notes to Table of Eddy-Stresses and Diffusivities.

I. 1917, June, 16d. 4h. 8m. L.A.T., hilltop near Ancemont, France. Standing hay composed of a species of *Festuca* (identified by my friend, Mr. SAM PIM). It grew fairly densely to 30 cm. from the ground and tall seed stems rose to 70 cm.

II. 1917, July, 16d. 19½h. L.A.T., Maffrecourt, France. Green oats 70 cm. high. No trees near. About sunset. Overcast with stratus. Observers: DAVID LONG and L. F. RICHARDSON.

III. 1917, June, 29d. 19h. 30m. L.A.T., Viel Dampierre, France. A field of corn 60 cm. high. Clouds (stratus) motionless. Observers: F. H. WEATHERALL, G. HUTCHINSON, L.F.R.

Eddy-stresses and of *Eddy-diffusivity* (continued).

$\widehat{yy}$ dynes per square centimetre.	$\widehat{hh}$ dynes per square centimetre.	$\widehat{xy}$ dynes per square centimetre.	$\widehat{yh}$ dynes per square centimetre.	$\widehat{hx}$ dynes per square centimetre.	$\frac{\partial \bar{v}_x}{\partial h}$ sec.$^{-1}$.	Conductivity. $= \rho K$ grm. cm.$^{-1}$ sec.$^{-1}$.	Diffusivity. K cm.^{2} sec.$^{-1}$.	Turbulivity. $\xi = g^2 \rho^2 K$ grm.^{2} cm.$^{-2}$ sec.$^{-5}$.
—	—	—	—	—	—	2·5	2,000	3,000
—	—	—	—	—	—	9·3	7,500	11,000
—	—	—	—	—	—	6	5,000	7,000
—	—	—	—	—	—	12	10,000	14,000
— —	-5 —							
—	—	—	—	—		120 —	10^5 —	$1 \cdot 4 \times 10^5$ —
—	—	—	—	—	—	$<0 \cdot 9 \times 10^3$	$<10^6$	$<0 \cdot 8 \times 10^6$
-110								
—	—	—	—	—	—	0·12	100	140
—	—	—	—	—	—	12	10,000	14,000
—	-1							

IV. 1917, June, 25d. 20h. 5m. L.A.T., Joinville, France. Moor with herbage dense to 10 cm. and rising thinly to 50 cm.

V. 1917, October, 4d. 8½h. G.M.T., Massiges, France. Flat moor with grass to 10 cm. and stems rising to 30 cm. Trees up-wind subtending an angle of 10 degrees. Overcast with strato nimbus, of which velocity/height = 0·025 sec.$^{-1}$. Temperature 0°·15 C. Observers: Olaf Stapledon and L. F. Richardson. Eddies partly due to observer.

VI. See photographs and description in this paper.

VII. 1917, July, 18d. 7¼h. L.A.T., Maffrecourt, France. Moor, with small trees, 5m. high and houses. Overcast.

VIII. 1917, August, 2d. 17h. L.A.T., South England. Overcast.

IX. 1917, July, 17d. 19h. 50m. L.A.T., East Champagne, France. Nine observations of a cloud at intervals of 1 minute by an Abney level. Angular elevation about 9 degrees.

X. 1917, August, 8d. 20h. L.A.T., Argonne, France. Large fire of petrol and wood. Smoke observed from distance of 10 km. with a sextant. If smoke were not hot, height of its upper edge above ground would have given $K = 4 \cdot 6 \times 10^5$ cm.^{2} sec.$^{-1}$, an over estimate. Irregularities in upper edge gave $K = 10^4$, probably an under estimate. Mean K of order of 10^5. Overcast.

XI 1918, April, 12d. 14h. 5m. L.A.T., France. Two anti-aircraft shell puffs at a mean elevation of 21 degrees above the horizon and 4 degrees apart in a vertical plane, were brought into coincidence in the field of view of a sextant. In 120 seconds their separation of 4 degrees did not vary visibly, certainly not by 4 minutes of arc. Their apparent motion was horizontal at $\frac{1}{1800}$ radian per sec. Height *assumed* 3200 metres—a likely value.

XII. 1914, November, 16d. 3h. to 9h. G.M.T. Taken from the re-production of the Dines anemogram on p. 81 of the 'Observer's Handbook,' Meteorological Office, London, 1917 edition.

XIII. 1917, July, 26d. 21h. L.A.T., the English Channel.

XIV. 1917, July, 26d. 16h. L.A.T., the English Channel, off Havre.

XV. 1917, July, 26d. 15h. L.A.T., the English Channel, off Havre. Observations of cloud at intervals of 1 minute with pocket sextant. Angular elevation about 4 degrees.

The photograph (2) also shows that the smoke spreads more rapidly upwards than downwards, indicating that the stress $\widehat{hh}$ and the turbulivity ξ both increase with height.

Fig. 3 shows another case of low eddy-conductivity occurring at sunset. The smoke here is from burning hydrogen phosphides; it is warm and rises slightly. To the eye the smoke appeared as a narrow wavy ribbon moving with a mean velocity of 1·2 metres per second. The broader smooth band shown in the photograph is due to the exposure of 75 seconds, made long in order to get an average effect. The source of smoke is a bottle 3·4 metres above ground and just within the picture. The bamboos are 5 metres apart. The air density was 0·00126 c.g.s. The eddy-diffusivity works out to about 130 c.g.s. units, the eddy-conductivity to 0·16 c.g.s., the turbulivity to 200 c.g.s. The sky was cloudless. Obstructions to windward rose above the horizon to an angle of only $\frac{1}{50}$ radian. The photograph was taken in latitude 51° 37′ N., longitude 4m. 24s. west, at 1919, Sept., 29d. 17h. 58m. G.M.T.

Above is a table of observations. It is noticeable that when two of the direct stresses $\widehat{xx}$, $\widehat{yy}$, $\widehat{hh}$ have been measured at the same time and place, they have been found to be not very unequal. G. I. TAYLOR has published some observations which show the same thing. It is as though there were a kind of equipartition of energy between the three components of the eddying motion. A very marked increase in both direct stress and diffusivity takes place either with velocity or with height. A rapid increase of viscosity with height in the first 200 metres has also been deduced by W. SCHMIDT from wind observations made by HELLMANN over a piece of flat land. ('Sitz. Akad. Wiss.,' Wien, 1917, Heft 6, p. 17).

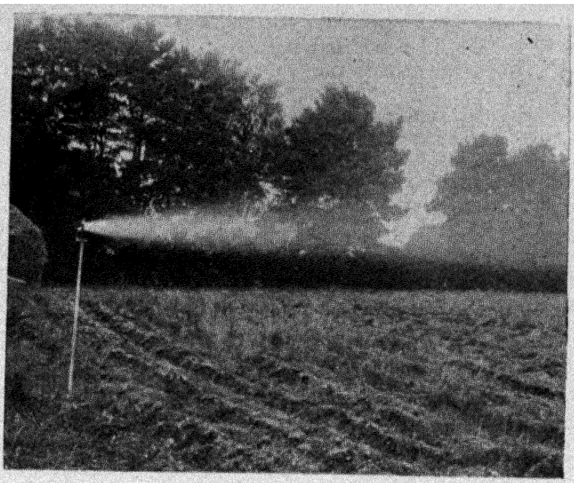


Fig. 1.

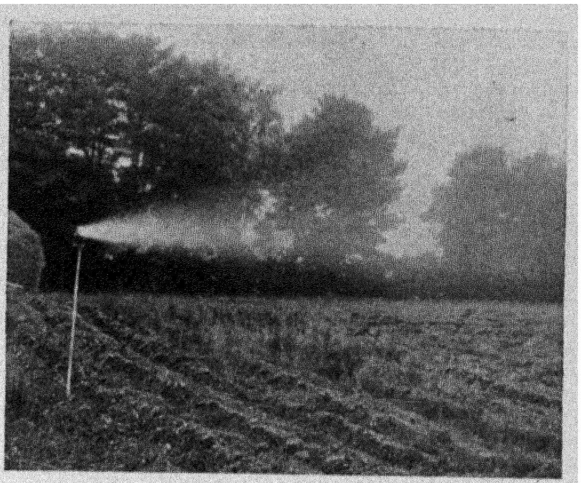
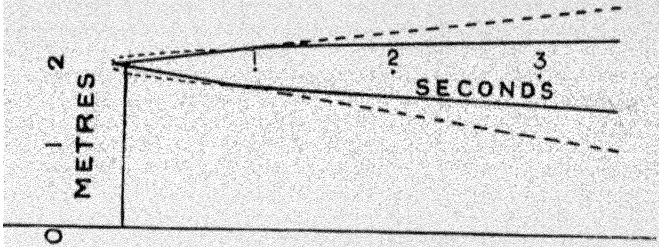
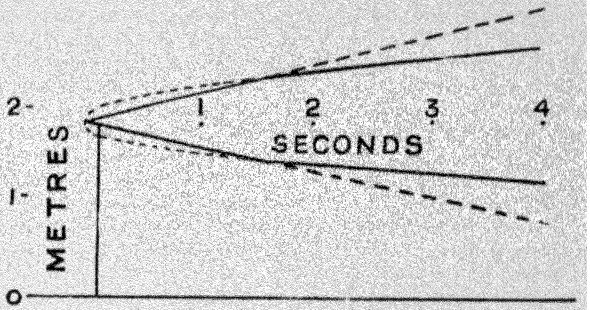


Fig 2.





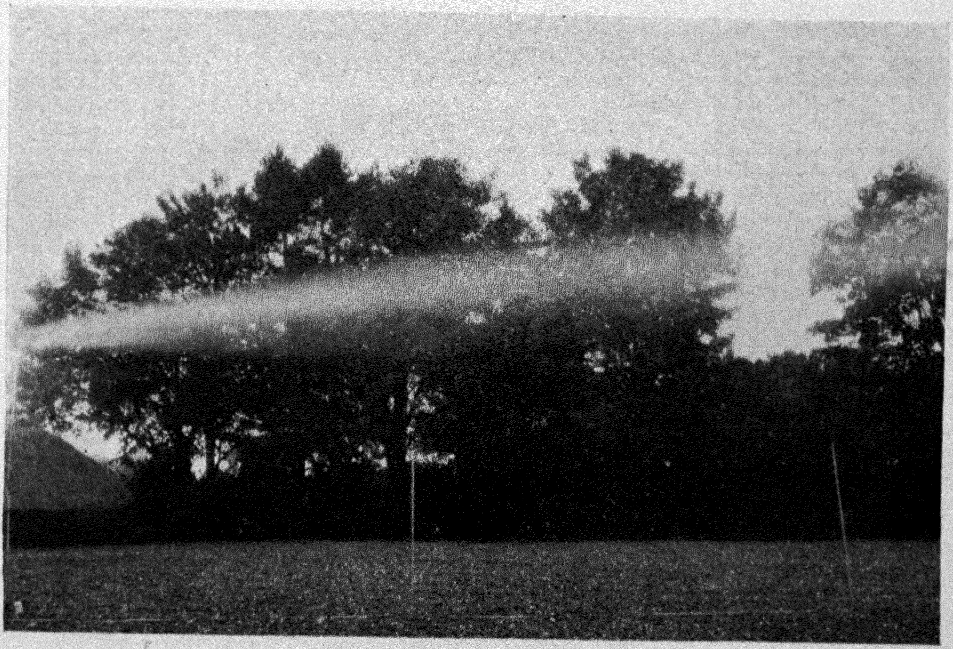


Fig. 3.

The six observations of steamers' smoke are put forward only as upper limits to the turbulence appropriate to the bare sea, for the steamer itself probably makes a considerable eddy.

It would be desirable to classify the observed eddy-conductivity as a function of four independent variables; namely, the height, the vertical gradient of entropy, the vertical gradient of velocity and the character of the surface. Vertical gradient of velocity is suggested as an independent because it measures the only rate-of-mean strain which attains a noticeable value in the free atmosphere, and because Osborne Reynolds[*] has shown that the energy of the eddy motion comes from the work done by the eddy-stresses upon the corresponding rates of mean strain. The observations here presented are much too scanty for such a classification, but to render the relation to height visible, the effect of velocity has been removed, in one sense, by dividing each value of the eddy-conductivity by the velocity at that level. The

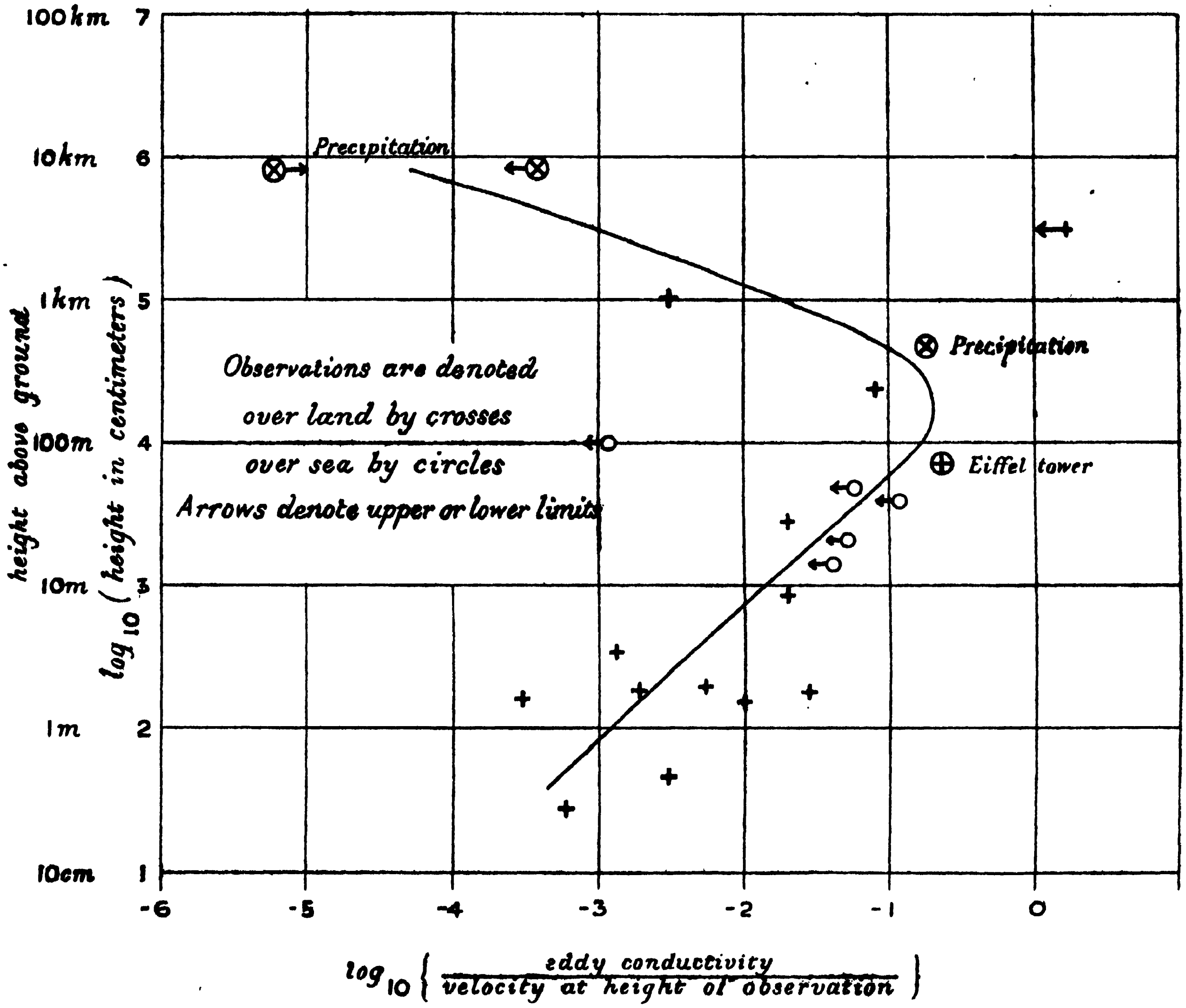


Fig. 4.

---

[*] Lamb, 'Hydrodynamics,' IV. edition, § 369, equation (21).

justification for this procedure is that TAYLOR has given reasons* for supposing that the viscosity, and therefore also the conductivity, is proportional to the velocity. For comparison with the present observations, the diagram shows TAYLOR'S mean value of the diffusivity at the Eiffel Tower,† and also some general means‡ deduced from precipitation by the writer.

In order to compare them it has been necessary to assume some corresponding velocities; for which purpose I have taken 540 cm. sec.^{-1} at the mean height of the Eiffel Tower, 700 cm. sec.^{-1} as a world-mean at 500 metres and 1000 cm. sec.^{-1} for the same at 8500 metres. These are based on information given in HANN'S 'Meteorology.' The conversion formulæ between eddy-conductivity, diffusivity and ξ have been given in Section I. In order to compress into a diagram the large ranges of height and conductivity, logarithms have been plotted. A smooth curve is drawn through the clustered observations over land. It shows a maximum between the heights of 100 and 1000 metres, and a marked falling off above and below. Not only c/v but also c the conductivity has a maximum here. TAYLOR'S first observations related to heights near this maximum and so he naturally came to the conclusion that there was no marked variation with height.

IX. CUMULUS EDDIES IN CALM WEATHER.

The familiar sequence, which can be observed in many places, is here illustrated by the mean of some selected days in latitude 49° in France, on the bare grass moors to the west of the forest of Argonne, in the month of May. The sun rose at 4h. 20m. local apparent time, but could not be seen for mist. By 6h. the disk of the sun became visible. At 7½h. the mist was rising in large pieces, leaving a brilliant blue sky. At 9h. the first cumuli appeared over the forest. About half-an-hour later they appeared over the grass land also. By noon the cumuli covered $\frac{4}{10}$ of the sky. By 16h. the cumuli had begun to spread out horizontally, and by 19h. they had vanished, leaving the sky clear again.

Now here we have a collection of eddies in which the rising parts, represented by the cumuli, visibly move to a level where they remain by mixing with their surroundings. So we should be able to calculate the diffusivity K by the direct application of the formula given by G. I. TAYLOR ('Phil. Trans.,' A, vol. 215, p. 3)

$$\mathrm{K} = \frac{1}{\mathrm{A}} \iint_{\mathrm{A}} v_{\mathrm{H}} (h-h')\, dx\, dy, \quad \ldots \ldots \ldots \quad (1)$$

where $h-h'$ is the height through which the air has moved before mixing, v_{H} is its vertical velocity, and A is a large horizontal area. Only, as TAYLOR'S formula assumes that ρ and K are independent of height, it may be as well to remove these restrictions. It is then found that ξ, as defined by equation (1), of Part I., is given by

---

* G. I. TAYLOR, 'Roy. Soc. Proc.,' A, vol. 92, pp. 196–199.

† G. I. TAYLOR, 'Roy. Soc. Proc.,' A, vol. 94 (1917), p. 141.

‡ L. F. RICHARDSON, 'Roy. Soc. Proc.,' A, vol. 96 (1919), p. 18.

$$\xi = \frac{g}{A} \iint_A m_H (p'-p)\, dx\, dy, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

where p' is the pressure at the initial level, m_H the vertical momentum per volume.

Now let us insert numerical values. The surface air was seen to begin to move at or a little before 7½h. The cumuli appeared to cease rising before 16h. and the height of their tops is known to average about 2 km. (*vide* HANN's 'Meteorology,' IIIrd edn., p. 280). Now if we suppose, as seems reasonable, that the top of the cumulus is formed from the damp air which was initially close to the ground, then the displacement, measured by pressure, is about 2 decibars, so that $(p'-p) = 2 \times 10^5$ dyne cm.$^{-2}$. The vertical velocity is 2 km. in 8·5 hours, that is 6·5 cm. sec.$^{-1}$. So the momentum per volume $= m \;= 7 \times 10^{-3}$ grm. cm.$^{-2}$ sec.$^{-1}$ in the rising current. The rising current covered 0·4 of the sky, so that averaging over the area A, as is done in (2), is equivalent to taking 0·4 of $m_H (p'-p)$ for the rising current. But the invisible descending currents contribute an equal amount to the integral. So

$$\begin{aligned} \xi &= g \times 0{\cdot}8 \times 2 \times 10^5 \times 7 \times 10^{-3} \\ &= 11 \times 10^5 \text{ grm.}^2 \text{ cm.}^{-2} \text{ sec.}^{-5}. \end{aligned}$$

This figure is about ten times greater than measures of ξ at a height of a few hundred meters, deduced by various authors. If the air which forms the top of the cumulus had really started from a height of 1 km. instead of from the ground, as we have supposed, then the numerical value of ξ would have to be divided by four.

Reasons have already been given (Part VII.) for supposing that ξ derived in this way from cumulus clouds is a measure of frictional effects, but not of the diffusion of entropy, because the linear term in (Part V., 3) does not vanish on taking the mean, owing to the fact that the eddies are produced by variations of entropy. To put it in another way: In G. I. TAYLOR's deduction of formula (1) the vertical gradient of the diffusing quantity is treated as not correlated with the vertical velocity. When we are dealing with cumulus clouds that assumption is probably justified if the diffusing quantity is horizontal velocity, but not if it is potential temperature.

To find ξ in the sense of diffusivity for potential temperature we should have to employ formula 32 of Part VII., namely

$$K = \frac{\overline{\left(\frac{\partial \theta}{\partial h}\right)' v'_H}}{\partial^2 \theta / \partial h^2},$$

For insertion in this we require lapse rates in cumulus clouds and in the clear air between them. Such have recently been obtained by airmen.

X. SUMMARY.

Part I. deals with notation. Measures of turbulence may advantageously be expressed in the form ξ in

$$\frac{\partial \chi}{\partial t} = \frac{\partial}{\partial p}\left(\xi \frac{\partial \chi}{\partial p}\right),$$

where p is the pressure (here used as a measure of height), t the time, and χ may be horizontal velocity in a fixed azimuth, or potential temperature, or water per mass of atmosphere. It is suggested that ξ might be called the "turbulivity." Its dimensions are grm.^{2} cm.$^{-2}$ sec.$^{-5}$. Better still is the conductivity $c = \xi \rho^{-1} g^{-2}$.

In Part II. the eddy-shearing stress on the ground is deduced from pilot balloon observations. Values on land in any self-consistent dynamical units are found to range from 0·0007 to 0·007 times the value of $\overline{m^2}/\rho$, where $\overline{m}$ is the mean momentum per volume up to a height of 2 km. and ρ is the density. Compare G. I. TAYLOR, 'Roy. Soc. Proc.,' A, vol. 92.

In Part III. evidence is given to show that the eddy-viscosity across the wind at Lindenberg increases with height, and, except near the ground, is much greater than the eddy-viscosity along the wind. Here ξ ranges from 10^4 to 5×10^5.

In Part IV. the spreading of a lamina of smoke is considered. Values of ξ ranging from 7 to 140,000 are found. ξ increases both with height and with velocity.

In Part V. the derivation of ξ from smoke observations is examined more thoroughly.

Part VI. deals with OSBORNE REYNOLDS' eddy-stresses. For one occasion an attempt was made to measure simultaneously all six components of stress by observing the motion of thistledown. The three direct stresses are easily measured. Not so the shearing stresses however, one was found to be 2·4 times its probable error.

Part VII. summarizes the theory of scattering of particles.

Part VIII. contains numerical values derived from scattering.

In Part IX. the turbulivity ξ is estimated from the rising of cumuli in calm weather and found to be 10^6, applicable only in the sense of friction. Thus the whole range of ξ observed in the free atmosphere was from 7 to a million in contrast with 0·2 in perfectly still air in a laboratory. The eddy-stresses observed have ranged in absolute value from 0·004 to 110 dynes cm.$^{-2}$.

II. *On a Theory of the Second Order Longitudinal Spherical Aberration for a Symmetrical Optical System.*

By T. Y. **Baker**, *B.A.*, *Instructor Commander*, *R.N.*, *and* L. N. G. **Filon**, *M.A.*, *D.Sc.*, *F.R.S.*, *Goldsmid Professor of Applied Mathematics and Mechanics in the University of London.*

Received December 2, 1919,—Read February 12, 1920.

§ 1. *Statement of the Problem and Historical References.*

If we consider a pencil of rays issuing from a point on the axis of a symmetrical optical system (*i.e.*, a system of refracting spherical surfaces, the centres of which lie on a straight line called the axis of the system), it is well known that, if the pencil be a thin one, of which the mean ray is along the axis, the first approximation to the emergent pencil is another punctual pencil, of which the rays pass through an image point, also situated on the axis. The general method of treatment of such image points, which are usually referred to as "geometrical" images, is due to **Gauss**, and is developed in any text book of Geometrical Optics.

When, however, the pencil considered is one of finite aperture, the outlying rays do not, after emergence, pass through the Gaussian image point, nor do they have the inclination assigned to them by the Gaussian calculation. The emergent rays lying in any one axial plane touch an envelope or caustic, which has one cusp at the Gaussian image, with the axis as proper tangent. The intercepts of any given emergent ray upon the axis and the image plane, measured from the Gaussian image, are known as the longitudinal and transverse spherical aberrations of that ray.

It is clear that if both these spherical aberrations, or either of them together with the inclination of the ray on emergence, be known for every possible position of object and image, and for every possible inclination of the incident ray, the whole complex of emergent rays lying in axial planes can be mapped out. The calculation of these aberrations is therefore of fundamental importance in practical optical design, where we do not deal with infinitely thin pencils.

The method employed hitherto for dealing with aberrations from the mathematical standpoint has been to develop the sines occurring in the refraction equations at each spherical surface in ascending power of some argument, which may be either the circular measure, or the sine, or the tangent, of one of the angles concerned, and to calculate, by the usual methods of successive approximation, the required aberrations as a series of ascending powers of such argument.

[Published June 16, 1920.

When this is done it is found that the terms due to the first power of the argument lead to the Gaussian image point, so that the series begin with a term involving the second power of the argument in the case of the longitudinal aberration, and the third power in the case of the transverse aberration. These terms are the first order. The following terms next in sequence, which are of fourth and fifth power respectively, are spoken of as aberrations of the second order, and so on.

A considerable amount of theoretical work has been done on aberrations of the first order by SEIDEL, ABBE and others, and the treatment of these is fairly well known. Unfortunately, it is found in practice that the first order aberrations do not give a sufficient approximation for the optician's requirements. In fact for a certain range of object and image positions, they are so badly out that they cannot be said to constitute an approximation at all. This fact has long been recognised by optical designers, whose practice is invariably to calculate, using the exact trigonometrical equations which involve no approximation at all, the correct paths of a number of selected rays, from which they draw conclusions as to the efficiency, or otherwise, of the proposed system from the practical point of view.

The trigonometrical method, however, from the designer's point of view, has the radical defect that, while it gives partial information about the performance of a given system, it gives no direct intimation of the direction in which the elimination of various defects is to be looked for, and it entails a long and laborious process of seeking for the optimum by trial and error.

The object of the authors of the present paper has been to develop a method of expressing the aberrations, which, while carrying the algebraic development to a stage including the second order, should be free from certain grave troubles involved by failure of convergency, troubles which appear to have been hitherto neglected. In fact this method gives numerical results that, for a single lens, are considerably more accurate than the ordinary second order formulæ. Further, these methods enable one to deal, in a comparatively easier form, with the problem of the second order aberrations of combinations of surfaces and systems, a problem which, so far as we know, has never been attacked from any general standpoint. KOENIG and VON ROHR (VON ROHR, 'Theorie der Optischen Instrumente,' Cap. V.) give a development of a formula for the coefficients of first order and second order in the longitudinal spherical aberration, based on ABBE'S method of Invariants, but so far as can be seen, no definite results are obtained for the second order terms.

DENNIS TAYLOR ('System of Applied Optics,' p. 67) gives a formula for the spherical aberration, developed in powers of the intercept made by the ray on the first principal plane, which includes terms of second order. But his formula, a particular case of those dealt with in the present paper, is limited to the thin lens, and no attempt seems to be made at anything like a general treatment of such aberrations.

Another important object of the method to be described is to express the aberrations in such a form that, in a combination of surfaces and lenses, the effect of a given surface or lens on the final result can be readily traced. This is fundamental for the designer, who usually proceeds to sketch out his system by Gaussian methods only, being guided therein by considerations of magnification, illumination, and field of view; and then goes on to eliminate the resulting image defects, so far as he can, by bending the lenses, *i.e.*, by altering their mean curvature without changing the focal length. In doing this he usually corrects one defect at a time, with the frequent result that, when, having corrected one defect by means of one lens, he proceeds to correct a second defect, he thereby causes the reappearance of the first.

If the effects of any given lens, however, are made apparent in the final formula, it becomes a more manageable problem to devise variations which will keep any one defect *invariant* whilst others are being dealt with.

§ 2. *Notation.*

There is no general agreement among mathematical writers as to the notation employed in dealing with optical problems, and it will be convenient to state here the symbols we have adopted. They are a modification of a system due to STEINHEIL.

The successive media, proceeding in the direction of travel of the light (from left to right in our figures), are denoted by even suffixes 0, 2, 4, &c., and the same suffixes affect the rays in these media, their inclinations, α_0, α_2, α_4, &c., to the axis, and their intersections I_0, I_2, I_4, &c., with that axis.

The successive geometrical images will be denoted by the letter J, thus J_0, J_2, J_4, &c. The successive surfaces of separation will be denoted by the odd suffixes 1, 3, 5, &c., and the same suffixes will affect the centres of curvature, the intersections of rays with the surfaces, and the points where the axis crosses the surfaces. The latter will be denoted by the letter A and the centres of curvature by the letter C.

Fig. 1 illustrates the use of this notation for two refracting surfaces.

The radii of curvature are r_1, r_3, r_5, &c., and are to be considered positive when $A_{2n+1}C_{2n+1}$ is measured from left to right.

The perpendicular from a centre of curvature on a ray is denoted by p and is affected by a double suffix, the first belonging to the centre of the curvature and the second to the ray. Thus p_{12} is the perpendicular from the centre of curvature C_1 of the first reflecting surface upon the ray in the second medium. Where there is no ambiguity the first suffix will usually be omitted.

The refractive index will be denoted by n and affected by the suffix of its medium.

Transverse magnifications will be denoted by M. The magnification produced by surface 1 will be denoted, as convenient, by M_1 or M_{02}; by surfaces 1, 3 combined either by M_{13} or M_{04}: by surfaces 1, 3, 5, combined either by M_{135}, or M_{06}, and so on.

The advantage of the double even suffix notation in this case is that we have a symbol, M_{40}, for the magnification when light passes backwards through the system, the order of the suffixes being material. Where odd suffixes are used, we have to use $1/M_1$, $1/M_{13}$, &c., for the reversed magnifications.

Ray magnifications will be denoted by $\mathbf{M}$. These are the limit of the sine-ratio for small inclinations, thus $\mathbf{M}_1 = \mathbf{M}_{02} = \underset{\alpha_0 \to 0}{\mathrm{L}} \sin \alpha_0 / \sin \alpha_2$. $M = \mathbf{M}$ when the initial and final media are the same.

With regard to inclinations, they will be treated as positive when the rays *converge* to the axis, as in fig. 1. The inclinations of the rays calculated by GAUSS' process will be denoted by β. Thus $\beta_0 = \alpha_0$, $\tan \beta_2 = \tan \alpha_0 / \mathbf{M}_{02}$, $\tan \beta_4 = \tan \alpha_4 / \mathbf{M}_{04}$, &c.

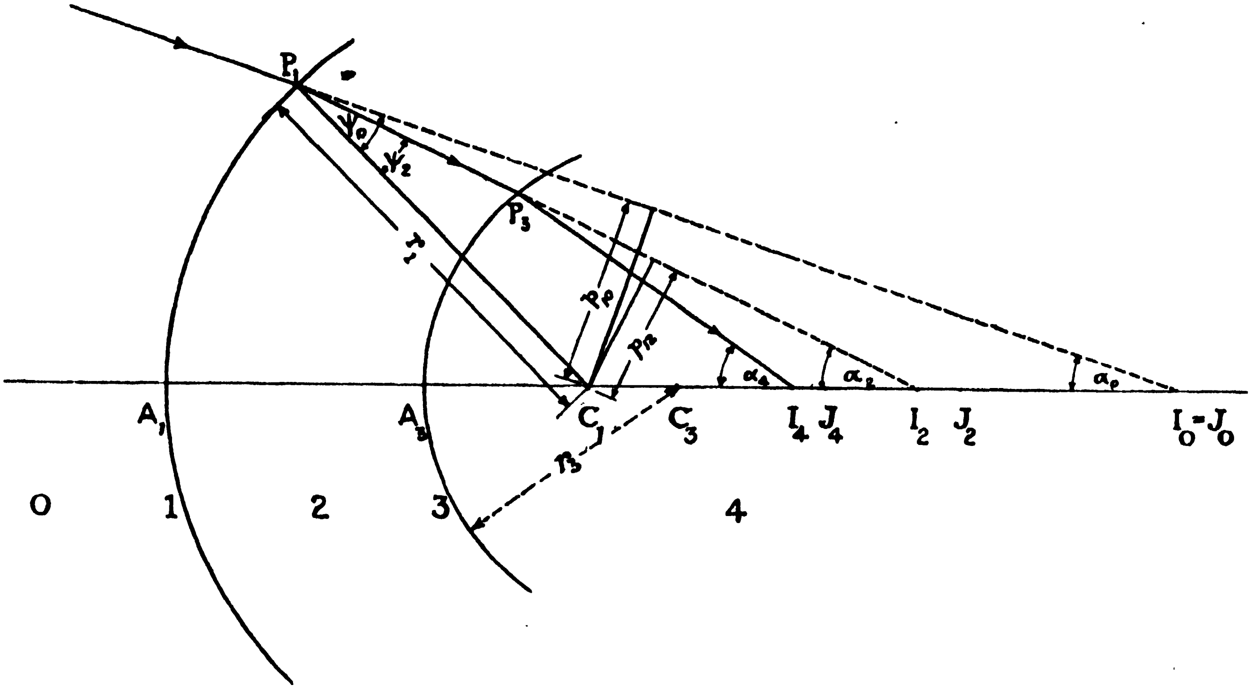


Fig. 1.

We may also use angles γ, calculated from a constant sine ratio, viz., $\gamma_0 = \alpha_0$, $\sin \gamma_2 = \sin \alpha_0 / \mathbf{M}_{02}$, $\sin \gamma_4 = \sin \alpha_0 / \mathbf{M}_{04}$, &c.

Throughout much of the work we shall use the same trigonometrical function (tangent or sine) of the angles α. If the tangent is used, we shall employ the following abbreviations:—

$$q_{2n} = \tan \alpha_{2n} \qquad t_{2n} = \tan \beta_{2n}.$$

If the sine is used, the meaning of q, t will be as follows:—

$$q_{2n} = \sin \alpha_{2n} \qquad t_{2n} = \sin \gamma_{2n}.$$

It will be found that many formulæ remain unaltered, whichever of the two interpretations for q and t is used.

All distances parallel to the axis will be denoted by x (the attribution of the symbol being indicated in each case) and will be measured positively from left to right.

The longitudinal aberration will be denoted by Δx, with the suffix of the medium, and will be reckoned positive when the point of intersection of the ray with the axis is to the right of the geometrical image. Thus $\Delta x_2 = J_2I_2$. This is opposite to the usual convention which is based on the fact that for positive or convergent lenses, I_2 is generally to the left of J_2; but, in the first place, this is not universally true, and, in the second place, the convention adopted by us was found more convenient in handling the algebra.

It is to be noted that, with the notation used, the well-known formula for a lens $\frac{1}{u}+\frac{1}{v}=\frac{1}{f}$ becomes $\frac{1}{v}-\frac{1}{u}=\frac{1}{f}$, the distances u and v being measured in the same direction.

The distances between successive refracting surfaces we denote by c, with the suffix of the medium.

In dealing with a system, especially where the initial and final media are not the same, it is very convenient to use an "equivalent" Gaussian system, in which lengths parallel to the axis are measured in each medium in terms of a unit proportional to its absolute refractive index.

If we denote the corresponding points in the equivalent Gaussian system by accents, we find that

$$A'_1J'_0 = \frac{A_1J_0}{n_0}, \qquad A'_1J'_2 = \frac{A_1J_2}{n_2}, \qquad A'_1A'_3 = c'_2 = \frac{c_2}{n_2},$$

$$A'_3J'_2 = \frac{A_3J_2}{n_2}, \qquad A'_3J'_4 = \frac{A_3J_4}{n_4}, \text{ \&c.}$$

If then we denote the quantities $\frac{n_2-n_0}{r_1}$, $\frac{n_4-n_2}{r_3}$, &c., by $\frac{1}{f_1}$, $\frac{1}{f_3}\ldots$; $f_1, f_3\ldots$ may be called the focal lengths of the successive refracting surfaces. The equations connecting image and object in the equivalent Gaussian system are

$$\frac{1}{A'_1J'_2}-\frac{1}{A'_1J'_0}=\frac{1}{f_1},$$

which is of the same form as the equation connecting image and object for a thin lens at A'_1.

Now bearing in mind that $A_3J_2 = A_1J_2 - A_1A_3$ and therefore $A'_3J'_2 = A'_1J'_2 - A'_1A'_3$, it is easy to show that the effects of the successive refracting surfaces in the actual system can be obtained by compounding a corresponding set of thin lenses in the equivalent Gaussian system. By dealing with the latter, we get rid of the asymmetry introduced by the difference of initial and final index. Of course this applies only to the calculation of the geometrical images. We note that in the equivalent Gaussian system the ray and transverse magnifications are identical and agree with the transverse magnifications in the actual system.

Finally the intercept of the incident ray on the leading principal plane of a system will usually be denoted by y. This is taken by some authors as the argument of the development in series, but differs only by a factor from tan α_0 or tan β_{2n}.

In many cases it will be convenient, in order to avoid unnecessarily large suffixes, to *condense* a system of surfaces or lenses, affecting quantities referring to the system itself with suffix 1, and the initial and final media with suffixes 0 and 2, the paths in the intermediate media not being explicitly considered.

§ 3. *Singularities and Convergency.*

Consider any symmetrical optical system, of which PL and QM (fig. 2) are the initial and final refracting surfaces. Let F_0 be the front focus of the system and UF_0V the caustic for backward-travelling rays which are parallel in the final medium.

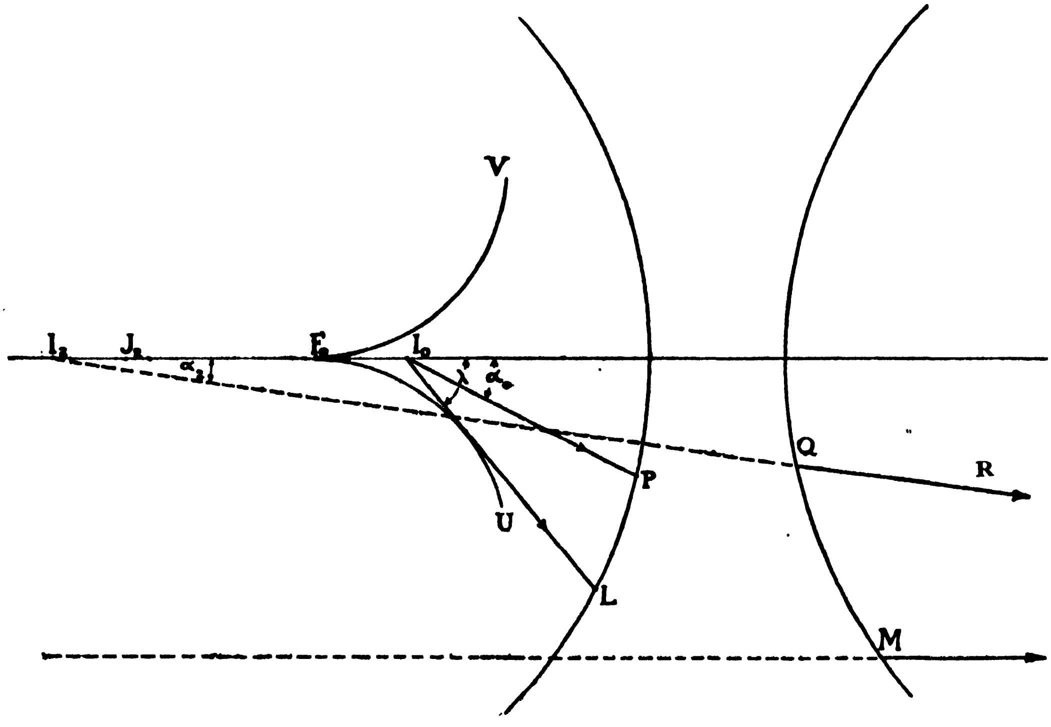


Fig. 2.

This caustic, as is well known, will usually be of the type shown in fig. 2, approximating to a semi-cubical parabola with a cusp at F_0, and, to fix ideas, we shall suppose the point of the cusp to be turned to the left. In the opposite case, an obvious modification of the argument will be found to lead to similar conclusions.

Any ray in the initial medium, which touches this caustic, must emerge parallel to the axis after passing through the system.

Let I_0 be an object point on the axis behind F_0 and sufficiently near to it for a real tangent to be drawn from I_0 to the caustic and yet go through the system. This will involve for I_0 a positive ray magnification M exceeding some definite *finite* limit. The Gaussian image point J_2 will then lie a finite distance in front of F_0.

Consider first a nearly paraxial incident ray I_0P. Such a ray will be refracted approximately according to the Gaussian law and will emerge at an inclination α_2, where α_2 is nearly equal to α_0/M and both α_0 and M being finite and positive, α_2 is also finite and positive. The ray emerges as QR, passing through a point I_2, finitely different from J_2.

As α_0 increases, α_2 at first increases with it, but as α_0 reaches the value λ, corresponding to the inclination of the ray I_0L which touches the front focus caustic, α_2 is again zero. Hence between those two values α_2 has at least one *maximum*, and for a given value of α_2 there are at least two values of α_0.

Thus, *within the range of values which are of practical importance*, α_0 is a many-valued function of α_2 having one or more branch-points, of which the one of least modulus corresponds to the first maximum of α_2.

Now, within the same range of values, all the aberrations must be given as single valued functions of α_0, since clearly there can only be one physical emergent ray, corresponding to one given physical incident ray. This statement, as we shall see, needs to be qualified when we are dealing with purely geometrical rays, but this need not affect the present stage of the discussion.

In consequence, if any aberration be expressed in terms of α_2—or of any trigonometrical function of α_2—that aberration must, in general, be a many-valued function of α_2, having for its branch-point of least modulus the first maximum value of α_2 mentioned above. It follows by a well-known result in theory of functions, that no TAYLOR'S series in α_2, or in sin α_2, or tan α_2, can be valid for values of α_2 exceeding this modulus numerically. For such values the series will be definitely divergent.

It is interesting to consider what happens when I_0 is on the other side of F_0, so that we are dealing with a large negative magnification. In this case no real tangent can be drawn from I_0 to the front focus caustic and the value of α_0, for which $\alpha_2 = 0$, is a pure imaginary. But here again, although we are now dealing with imaginary values, we get two values of α_0 for a given (pure imaginary) value of α_2, and, although no such maximum of α_0 occurs in the purely real values, the modulus of the imaginary branch-point limits the validity of TAYLOR'S series in α_2 as before.

Thus there exists always a certain range, extending a finite distance (depending on the nature of the optical system) on either side of the front focus, within which no development of any aberration in powers of α_2 or of its trigonometrical functions (or, indeed, by similar reasoning, of any inclination of the ray, except in the original medium) is valid for the whole pencil of rays which actually traverse the system.

Indeed, as the object point I_0 approaches the front focus, it is clear that both λ and the maximum α_2 tend to zero, so that only an infinitesimal portion of the rays can be dealt with by the method of successive aberrations, *i.e.*, by the TAYLOR'S series.

That the range of failure is by no means an unimportant one is shown by an example given by the authors in a paper read before the Optical Society in December, 1918. In this example the system considered is a positive lens of unit focal length and thickness $\frac{1}{16}$, meniscus shaped, with curvatures 1 and 2·36, and its convex side towards the incoming light. For such a lens and magnification as low as 2, the critical value of α_4 is found to be about 4° 40′, corresponding to a value of α_0 of 13°, whilst the greatest practical value of α_0 is 26°, so that in this case only about $\frac{1}{4}$ of the light going through the lens could be dealt with by series in terms of the emergent angle. From $M = 2$ to $M = \infty$ the conditions are still worse.

As a matter of fact, it appears that in this case the range of magnifications, within which development in terms of the emergent inclinations is possible for all rays travelling through the lens, is restricted to a range lying somewhere between $M = -1$ and $M = 1·5$. This makes it clear that we cannot depend, in the calculation of the aberrations of an optical system, upon any series with the emergent inclination as argument. This is important, because from other considerations it would have been valuable to have been able to express the equation of the emergent ray in the form

$$y + qx = f(q)$$

where q is the inclination of the emergent ray, and to proceed to obtain successive approximations to the caustic by developing $f(q)$ in powers. It now appears that this is not, in general, legitimate.

We now come to the consideration of series proceeding by powers of α_0, or of its trigonometrical functions. Here the question of many-valuedness will not occur, except as follows.

If we consider a ray impinging upon a spherical refracting surface, this ray, if produced, will meet the surface at a second point. Treating the problem from the purely analytical standpoint, this second point is also one at which refraction takes place, and thus, for the same α_0, there will, in general, be two values of α_2, four of α_4, and so on. α_{2n} will therefore, in general, be a multiple-valued function of α_0, and the aberrations will also be multiple-valued functions, and the branch-points of these multiple-valued functions will, as before, limit the convergency of the TAYLOR series.

Now clearly two branches coincide whenever there occurs a grazing incidence; and, therefore, if the system be so arranged (as it almost necessarily is) so that no grazing incidence is reached, there will be no *real* branch-points within the range of practical values. But this does not mean that the TAYLOR'S series will necessarily be valid, for there might be imaginary branch-points. A very simple example will show how such branch-points can occur.

If I_0 be a source of light placed in front of a plate of thickness c_2 and refractive index n, the perpendicular from I_0 on the plate being the axis of the system, it is easily verified that the longitudinal spherical aberration

$$J_4I_4 = c_2\left(\frac{1}{n} - \frac{\tan\alpha_2}{\tan\alpha_0}\right),$$

where $\sin\alpha_0 = n\sin\alpha_2$, so that

$$J_4I_4 = \frac{c_2}{n}\left(1 - \frac{\sqrt{1-\sin^2\alpha_0}}{\sqrt{1-\frac{\sin^2\alpha_0}{n^2}}}\right).$$

The branch-points here correspond to $\alpha_0 = \frac{1}{2}\pi$ or $\alpha_2 = \frac{1}{2}\pi$, *i.e.*, to grazing incidence at the first or second surface respectively.

Clearly if $n > 1$, then, since $\sin^2\alpha_0 \leq 1$, the second grazing incidence can never occur for real values of α_0.

But if we take as our argument $t_0 = \tan\alpha_0$, which removes the first branch-point to infinity, we find

$$J_4I_4 = \frac{c_2}{n}\left(1 - \frac{1}{\sqrt{1+t_0^2\left(1-\frac{1}{n^2}\right)}}\right),$$

and this has imaginary branch-points where $t_0 = \pm\frac{in}{\sqrt{(n^2-1)}}$. The radius of convergence of the TAYLOR'S series in t_0 is therefore given by $t_0 = \frac{n}{\sqrt{(n^2-1)}}$, a value which does not correspond to any physical limitation of the rays. This applies to both the longitudinal and the transverse spherical aberrations in this case.

The above example also brings out another important point; for if in it $\sin\alpha_0$ is taken as the argument, the branch-points are ± 1, $\pm n$; both of which correspond to definite physical limitations, viz., grazing incidence and total internal reflection, so that in this case the limitations of the TAYLOR'S series are also the limitations of the problem.

We see then that the validity even of the expansion in α_0 may be limited by the existence of branch-points, and that the choice of the particular trigonometrical function in which we expand may exercise a considerable influence on the result.

The limitation of the α_0 developments due to branch-points will not, however, as in the case of the α_2 developments, lead to vanishing radii of convergence. There is always a finite region within which these developments may be used. In what follows, therefore, we have exclusively used α_0 as argument.

In dealing with the longitudinal spherical aberration another limitation presents itself. We have seen that if $\alpha_0 = \lambda$ (fig. 2), $\alpha_2 = 0$. It follows that the intersection of the emergent ray with the axis is then at infinity, or the longitudinal aberration is infinite and afterwards changes sign. Thus the values $\alpha_0 = \pm\lambda$ correspond to poles of the longitudinal aberration. These poles, being the singularities of least modulus, govern the convergence of the TAYLOR'S series in this neighbourhood, and this radius of convergence tends to zero as the object approaches the front focus. It was a consideration of this difficulty which primarily led us to put the longitudinal aberration in a new form.

This difficulty does not arise with the transverse spherical aberration. The poles of the longitudinal spherical aberration are due to the zeros of α_0, and on multiplying by $\frac{\tan\alpha_0}{\mathbf{M}}$ to get the transverse aberration, these poles disappear.

§ 4. *Summary of Method and Results.*

The general principle of the method employed was suggested by an attempt to fit an empirical formula to the longitudinal aberration of a lens for a certain range of curvatures and object and image positions. This empirical formula was discussed by the authors in a paper recently read before the Optical Society* and was found to give, on the whole, a singularly good fit. Briefly stated, the formula is of the following type :—

$$\Delta x = \frac{\mathrm{A}t^2}{1+\mathrm{B}t^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where t is the slope of the emergent "Gaussian" ray, so that $t = t_0/\mathbf{M}$. A is the (known) theoretical constant of the first-order aberration, which is a quartic in the magnification, and B is a cubic in the magnification, the coefficients in which are determined empirically. This formula was found to give a good approximation, even when the magnification was high and we were working well outside the limits of convergency of the TAYLOR'S series for Δx.

If we consider any given object point, the longitudinal spherical aberration will be a function of t_0, that is, of t. Denoting it by $f(t)$, the reasoning of the preceding section shows that $f(t)$ is always one-valued for a finite (and generally quite considerable) range of t, but it is not regular, having poles at $t = \pm\,\tau$, where $\tau = \tan\lambda/\mathbf{M}$ and becomes rapidly small as the magnification increases numerically.

If, however, we write $(1-t^2/\tau^2)\,f(t) \equiv \phi(t)$, $\phi(t)$ is now limited only by the original branch-points of $f(t)$ and will, in general, have an adequate radius of convergence. We may therefore expand it in a TAYLOR'S series, and we get for $f(t)$ the form

$$\Delta x = f(t) = \frac{at^2+bt^4+ct+\ldots}{1-t^2/\tau^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

---

* BAKER and FILON, "On an Empirical Formula for the Longitudinal Spherical Aberrations in a Thick Lens," 'Proceedings of the Optical Society,' December, 1918.

If the series in the numerator converges rapidly, it will be sufficient, provided a is not near zero, to stop at the first term, and we get as an approximate formula

$$\Delta x = \frac{at^2}{1-t^2/\tau^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (3)$$

which is of the same form as (1).

We have necessarily $a = A$, and if the two formulæ are to tally we should have in addition $B = -1/\tau^2$.

The formula in the form (3), however, is not rigorously correct to the second order of aberrations inclusive, unless b happens to be small. If we wish to retain second order terms complete, we have to use

$$\Delta x = \frac{at^2+bt^4}{1-t^2/\tau^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

and this can be written, to the same order of *algebraic* approximation, in the form

$$\Delta x = at^2/\{1-(1/\tau^2+b/a)\,t^2\}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

provided again a is not zero.

If this form (5) is adopted, then the B of the empirical formula should be $1/\tau^2+b/a$. But if this is done, the formula suffers from two defects: (i) it fails whenever a is near zero; (ii) it does not give exact compensation for the poles in the critical range for M large.

The further question then arose: how far are formulæ of type (4) or (5) suitable for dealing with *combinations* of surfaces or lenses? An important guiding consideration, in all work of this kind, must be the relative simplicity of the formulæ in passing from a single surface or lens to a combination, and whether these formulæ are suitable for tracing the effect of individual surfaces or lenses upon the final result.

We have ultimately been led to the conclusion that no *single* formula can satisfy completely the three ideal requirements, viz.: (i) exact agreement with development as far as the second order inclusive; (ii) simplicity in dealing with combinations; (iii) exact compensation of the poles in the critical range of M.

The method finally adopted satisfies conditions (i) and (ii). It only satisfies (iii) approximately. Numerical calculations show that numerically the approximation is adequate in the case of a lens or a simple surface. In the case of more complicated systems we have, as yet, no numerical data.

The first part of the investigation deals with the single refraction. It is there shown that the longitudinal aberration can be put into the form (1), *i.e.*, $\Delta x = At^2/(1+Bt^2)$, where the formula is correct to the second order inclusive.

We also find, for the inclination of the emergent ray, the formula

$$q = t\,(1+Bt^2)/(1+Ct^2), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (6)$$

which is correct to the first order when q and t are tangents and to the second order when q and t are sines.

In the above A, B, C are polynomials in M of degrees 4, 3, 2 respectively, so that the empirical formula is well justified for the simple refracting surface. In this case, too, it is possible to calculate τ directly, and, in fact, a simple geometrical construction is given for it. When this is followed for varying image-positions, it is found that outside a certain range of M, the τ so obtained becomes irrelevant, and that, in fact, if the correct factor $1-t^2/\tau^2$ is retained in the denominator of Δx, although it improves the fit by removing singularities in the range round $M = \infty$, it introduces entirely fictitious singularities in other and important parts of the range, and makes the formula worthless.

A good deal of light is thrown upon the problem when it is found that, if we develop $\frac{1}{\tau^2}$ in descending powers of M in the neighbourhood of $M = \infty$, the two leading terms are discovered to be identical with the two leading terms of the cubic B, previously obtained. This makes our B approximate more and more closely to $\frac{1}{\tau^2}$ precisely as the effect of the denominator term becomes more important, and it is this fact which is the key to the numerical value of the method.

We then proceed to show how the constants for a combination of the two systems can be obtained from the corresponding constants of the individual systems. In doing this it appears that, so soon as we pass from the single refracting surface to the lens, a new constant is introduced into the formula, which now takes the form

$$\Delta x = \frac{At^2 + Et^4}{1 + Bt^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where A and B are of the same form as before, but E is now a polynomial of degree 6 in M. In the case of a lens the term Et^4 is found to be, in general, of small importance, which accounts for the good fit of the empirical formula.

The formula (7) *for a combination* holds good to the second order inclusive, and B agrees with $\frac{1}{\tau^2}$, when M is large, as far as the leading term only. For numerical purposes, however, a correction is discussed, which is very readily applied, and which makes the two leading terms in B agree with the two leading terms in $\frac{1}{\tau^2}$, as in the case of the single refracting surfaces.

The formulæ for combining two systems take comparatively simple forms; the A, B and C for the combination are expressed as linear functions of the A's, B's and C's of the components, and the E as a lineo-linear function of the A's, B's and C's of the components, each term involving a product of which one factor belongs to one component of the combination and the other factor to the second. In addition the E for the combination involves linear terms in the E's of the components.

These results are found to hold good in the more general case of the combination of three or more systems. It will follow that if the constants A, B, C, E are tabulated for lenses of various curvatures, the effect of the combination can be traced relatively easily and the aberrations corrected, so far as possible, by suitably bending the lenses, while keeping the general arrangement and the magnifications the same.

Explicit values of the constants for the single refracting surface and a single thick or thin lens have been obtained and are tabulated for reference, so as to be available for eventual computation of the required tables. We have also given some numerical values for a single lens, and a numerical test of the accuracy in this case, which works out at about $\frac{1}{500}$ of the total aberration for the range of cases taken.

The corresponding formulæ with sin γ instead of tan β as argument are discussed, and it is shown that the equations of combination are of the same form as before.

Certain invariant relations between the coefficients in A, B, C, E are developed, which enable various calculations to be simplified and in particular to determine these constants for a system reversed, when they are known for the direct system. This will generally halve the work of tabulation.

§ 5. *The Single Refracting Surface.*

Using the general notation described in § 2, consider refractions at a single refracting surface.

Let ψ_0, ψ_2 denote the angles of incidence and refraction, so that $\psi_0 = C_1P_1I_0$, $\psi_2 = C_1P_1I_2$ (fig. 1).

Let $C_1I_0 = X_0 = x_0$, $C_1I_2 = X_2 = x_2 + \Delta x_2$, $C_1J_2 = x_2$ we then have the set of refraction equations

$$\sin \psi_0 = p_0/r_1 = x_0 \sin \alpha_0/r_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

$$\sin \psi_2 = p_2/r_1 = X_2 \sin \alpha_2/r_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

$$n_2 p_2 = n_0 p_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

$$\alpha_2 - \alpha_0 = \psi_0 - \psi_2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Let $\xi_0 = A'_1I'_0$, $\xi_2 = A'_1J'_2$ in the "equivalent" Gaussian system (see § 2). Then

$$n_0\xi_0 = x_0 + r_1, \qquad n_2\xi_2 = x_2 + r_1, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

and we find, using the first approximation when α_0, &c., are small

$$1/\xi_2 - 1/\xi_0 = 1/f_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

If $\mathbf{M}_1 = \underset{\alpha_0 \to 0}{\mathrm{L}}\, \alpha_0/\alpha_2$, which we shall call the *ray magnification*, we have

$$\mathbf{M}_1 = n_2\xi_2/n_0\xi_0 = n_2x_2/n_0x_0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (14)$$

The transverse magnification M_1 is given by

$$\mathrm{M}_1 = \xi_2/\xi_0 = x_2/x_0 = n_0\mathbf{M}_1/n_2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (15)$$

We can also express (13) in another well-known form, namely,

$$n_0/x_2 - n_2/x_0 = 1/f_1, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (16)$$

whence, using (14), we obtain

$$\left.\begin{aligned} x_0 &= n_2 f_1 (1-\mathbf{M}_1)/\mathbf{M}_1 \\ x_2 &= n_0 f_1 (1-\mathbf{M}_1) \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (17)$$

Again, from (9), (8) and (10)

$$\begin{aligned} \mathrm{X}_2 &= p_2/\sin\alpha_2 = n_0x_0\sin\alpha_0/n_2\sin\alpha_2 = n_0f_1(1-\mathbf{M}_1)\sin\alpha_0/\mathbf{M}_1\sin\alpha_2 \\ &= x_2\sin\alpha_0/\mathbf{M}_1\sin\alpha_2, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (18) \end{aligned}$$

and the longitudinal aberration

$$\Delta x_2 = \mathrm{X}_2 - x_2 = x_2(\sin\alpha_0/\mathbf{M}_1\sin\alpha_2 - 1). \quad . \quad . \quad . \quad . \quad . \quad . \qquad (19)$$

The corresponding longitudinal aberration in the equivalent Gaussian system is found from

$$\Delta\xi_2 = \Delta x_2/n_2 = (n_0 f_1/n_2)(1-\mathbf{M}_1)(\sin\alpha_0/\mathbf{M}_1\sin\alpha_2 - 1). \quad . \quad . \quad . \quad . \qquad (20)$$

Now from (11)

$$\begin{aligned} \sin\alpha_2 &= \sin\alpha_0\cos\psi_0\cos\psi_2 + \sin\psi_0\cos\alpha_0\cos\psi_2 - \sin\psi_2\cos\alpha_0\cos\psi_0 \\ &\quad + \sin\alpha_0\sin\psi_0\sin\psi_2, \end{aligned}$$

whence, using (8), (9) and (10),

$$\begin{aligned} \sin\alpha_2/\sin\alpha_0 = {} & \{1-(x_0/r_1)^2\sin^2\alpha_0\}^{\frac{1}{2}}\{1-(n_0x_0/n_2r_1)^2\sin^2\alpha_0\}^{\frac{1}{2}} \\ & + (x_0/r_1)\{1-\sin^2\alpha_0\}^{\frac{1}{2}}\{1-(n_0x_0/n_2r_1)^2\sin^2\alpha_0\}^{\frac{1}{2}} \\ & - (n_0x_0/n_2r_1)\{1-\sin^2\alpha_0\}^{\frac{1}{2}}\{1-(x_0/r_1)^2\sin^2\alpha_0\}^{\frac{1}{2}} \\ & + (n_0x_0^2/n_2r_1^2)\sin^2\alpha_0, \end{aligned}$$

and developing this in ascending powers of $\sin\alpha_0$, we obtain, retaining only terms of fourth degree

$$\begin{aligned} \sin\alpha_2/\sin\alpha_0 = {} & 1 + x_0(n_2-n_0)/n_2r_1 - \tfrac{1}{2}\mathrm{P}\sin^2\alpha_0 \\ & - \tfrac{1}{8}\mathrm{P}\sin^4\alpha_0\{(n_2+n_0)^2x_0^2 + n_2(r_1-x_0)(n_2r_1+n_0x_0)\}/n_2^2r_1^2, \end{aligned}$$

where

$$\mathrm{P} = (1-n_0/n_2)(x_0/r_1)(1+x_0/r_1)(1-n_0x_0/n_2r_1),$$

and is the quantity whose vanishing gives the aplanatic points and must therefore be a factor of every coefficient after the first in the development of $\sin\alpha_2$ in powers of $\sin\alpha_0$.

If we write for shortness

$$Q \equiv (1+n_0/n_2)^2 x_0^2/r_1^2 + (1-x_0/r_1)(1+n_0x_0/n_2r_1),$$

and remember that

$$1+x_0(n_2-n_0)/n_2r_1 = 1/\mathbf{M}_1,$$

we find

$$\begin{aligned}\sin\alpha_2/\sin\alpha_0 &= 1/\mathbf{M}_1 - \tfrac{1}{2}\mathrm{P}\sin^2\alpha_0/(1-\tfrac{1}{4}\mathrm{Q}\sin^2\alpha_0)\\ &= \mathbf{M}_1^{-1}\{1-(\tfrac{1}{2}\mathrm{P}\mathbf{M}_1+\tfrac{1}{4}\mathrm{Q})\sin^2\alpha_0\}/(1-\tfrac{1}{4}\mathrm{Q}\sin^2\alpha_0)\\ &= \mathbf{M}_1^{-1}\{1+\mathbf{B}\sin^2\alpha_0/\mathbf{M}_1^2\}/\{1+\mathbf{C}\sin^2\alpha_0/\mathbf{M}_1^2\} \quad . \quad . \quad (21)\end{aligned}$$

correct as far as the second order inclusive, where

$$\mathbf{B} = -\tfrac{1}{2}\mathrm{P}\mathbf{M}_1^3 - \tfrac{1}{4}\mathrm{Q}\mathbf{M}_1^2, \qquad \mathbf{C} = -\tfrac{1}{4}\mathrm{Q}\mathbf{M}_1^2,$$

from which, after some reductions

$$\mathbf{C} = -\tfrac{1}{4}(n_2-n_0)^{-2}\{(n_2^2+n_0n_2+n_0^2)-3(n_0^2+n_2^2)\mathbf{M}_1+3(n_2^2-n_0n_2+n_0^2)\mathbf{M}_1^2\}, \quad (22)$$

and

$$\begin{aligned}\mathbf{B} &= \tfrac{1}{2}(n_2-n_0)^{-2}(1-\mathbf{M}_1)(n_2-n_0\mathbf{M}_1)(n_0-n_2\mathbf{M}_1)+\mathbf{C}\\ &= \tfrac{1}{4}(n_2-n_0)^{-2}\{-(n_2^2-n_0n_2+n_0^2)+(n_2-n_0)^2\mathbf{M}_1-(n_2^2+n_0^2-5n_0n_2)\mathbf{M}_1^2-2n_0n_2\mathbf{M}_1^3\}. (23)\end{aligned}$$

Returning to equation (19) and using (21)

$$\begin{aligned}\Delta x_2 &= x_2(\mathbf{C}-\mathbf{B})\mathbf{M}_1^{-2}\sin^2\alpha_0/\{1+\mathbf{B}\mathbf{M}_1^{-2}\sin^2\alpha_0\}\\ &= n_2f_1\mathrm{A}\mathbf{M}_1^{-2}\sin^2\alpha_0/\{1+\mathbf{B}\mathbf{M}_1^{-2}\sin^2\alpha_0\}, \quad . \quad . \quad . \quad . \quad . \quad (24)\end{aligned}$$

where

$$\begin{aligned}\mathrm{A} &= x_2(\mathbf{C}-\mathbf{B})/n_2f_1\\ &= -\tfrac{1}{2}(n_2-n_0)^{-2}(n_0/n_2)(1-\mathbf{M}_1)^2(n_2-n_0\mathbf{M}_1)(n_0-n_2\mathbf{M}_1)\\ &= \tfrac{1}{2}(n_2-n_0)^{-2}(n_0/n_2)\left\{\begin{array}{l}-n_0n_2+(n_2+n_0)^2\mathbf{M}_1-2(n_2^2+n_0n_2+n_0^2)\mathbf{M}_1^2\\ \quad +(n_2+n_0)^2\mathbf{M}_1^3-n_0n_2\mathbf{M}_1^4\end{array}\right\}. \quad (25)\end{aligned}$$

All the above formulæ are correct to the second order of aberrations inclusive. We note that A, **B** and **C** are polynomials of degree 4, 3 and 2 in the magnification respectively.

If we express the aberration in terms of tangents instead of sines we have at once

$$\begin{aligned}\Delta x_2 &= n_2f_1\mathrm{A}\mathbf{M}_1^{-2}\tan^2\alpha_0/\{1+(\mathbf{B}+\mathbf{M}_1^2)\mathbf{M}_1^{-2}\tan^2\alpha_0\}\\ &= n_2f_1\mathrm{A}\tan^2\beta_2/\{1+\mathrm{B}\tan^2\beta_2\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)\end{aligned}$$

where

$$
\begin{aligned}
\mathrm{B} &= \mathbf{B}+\mathbf{M}_1^2 \\
&= \tfrac{1}{4}(n_2-n_0)^{-2}\{-(n_2^2-n_0n_2+n_0^2)+(n_2-n_0)^2\mathbf{M}_1+3(n_2^2-n_0n_2+n_0^2)\mathbf{M}_1^2-2n_0n_2\mathbf{M}_1^3\}. \quad (27)
\end{aligned}
$$

When, however, we come to develop a formula for tangents, similar to (21) for sines, it is found that, in the series for $\tan\alpha_2/\tan\alpha_0$, we do not have all the coefficients after the first vanishing together; for, even at the aplanatic points, the tangent ratio is not constant.

We can, indeed, write

$$
\tan\alpha_2/\tan\alpha_0 = \mathbf{M}_1^{-1}(1+\mathrm{B}\mathbf{M}_1^{-2}\tan^2\alpha_0)/(1+\mathrm{C}\mathbf{M}_1^{-2}\tan^2\alpha_0), \quad . \quad . \quad . \quad (28)
$$

and choose the coefficients B and C so that the developments shall agree as far as terms in $\tan^4\alpha_0$ inclusive. This can, in general, be done in one way only. But we then find that B and C are no longer integral functions of the magnification. Their infinities have to be taken into account, and generally the method becomes complicated and unsatisfactory.

From other considerations, however, it appears that since the zeroes of α_2 must be the same as the poles of Δx_2, the B in (28) must be the same as the B in (26), and this will fix C as follows :—

$$
\begin{aligned}
\tan^2\alpha_2 &= \sin^2\alpha_2/(1-\sin^2\alpha_2) \\
&= \frac{\mathbf{M}_1^{-2}\sin^2\alpha_0\{1+\mathbf{B}\sin^2\alpha_0/\mathbf{M}_1^2\}^2}{\{1+\mathbf{C}\sin^2\alpha_0/\mathbf{M}_1^2\}^2-\mathbf{M}_1^{-2}\sin^2\alpha_0\{1+\mathbf{B}\sin^2\alpha_0/\mathbf{M}_1^2\}^2} \\
&= \frac{\mathbf{M}_1^{-2}\tan^2\alpha_0\{1+(\mathbf{B}+\mathbf{M}_1^2)\mathbf{M}_1^{-2}\tan^2\alpha_0\}^2}{(1+\tan^2\alpha_0)\{1+(\mathbf{C}+\mathbf{M}_1^2)\mathbf{M}_1^{-2}\tan^2\alpha_0\}^2-\mathbf{M}_1^{-2}\tan^2\alpha_0\{1+(\mathbf{B}+\mathbf{M}_1^2)\mathbf{M}_1^{-2}\tan^2\alpha_0\}^2},
\end{aligned}
$$

substituting from (21).

Hence, taking the square root and developing the denominator in powers of $\tan\beta_2$, *i.e.*, of $\mathbf{M}_1^{-1}\tan\alpha_0$,

$$
\begin{aligned}
\tan\alpha_2/\tan\beta_2 &= \frac{(1+\mathrm{B}\tan^2\beta_2)}{1+\tan^2\beta_2(\mathbf{C}+\tfrac{3}{2}\mathbf{M}_1^2-\tfrac{1}{2})+\tfrac{1}{2}\tan^4\beta_2\{\mathbf{C}(1+\mathbf{M}_1^2)+\tfrac{3}{4}\mathbf{M}_1^4+\tfrac{3}{2}\mathbf{M}_1^2-\tfrac{1}{4}-2\mathrm{B}\}} \\
&= (1+\mathrm{B}\tan^2\beta_2)/(1+\mathrm{C}\tan^2\beta_2+\mathrm{D}\tan^4\beta_2), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)
\end{aligned}
$$

where

$$
\left.
\begin{aligned}
\mathrm{C} &= \mathbf{C}+\tfrac{3}{2}\mathbf{M}_1^2-\tfrac{1}{2} \\
\mathrm{D} &= \mathbf{C}(1+\mathbf{M}_1^2)+\tfrac{3}{4}\mathbf{M}_1^4+\tfrac{3}{2}\mathbf{M}_1^2-\tfrac{1}{4}-2\mathrm{B} \\
&= \mathrm{C}(1+\mathbf{M}_1^2)+\tfrac{1}{4}(1-\mathbf{M}_1^2)(1+3\mathbf{M}_1^2)-2\mathrm{B}
\end{aligned}
\right\}, \quad . \quad . \quad . \quad . \quad . \quad (30)
$$

(29) is now correct as far as the second order inclusive.

If we only require the tangent ratio correct up to the first order of aberrations, we have the formula

$$
\tan\alpha_2 = \tan\beta_2(1+\mathrm{B}\tan^2\beta_2)/(1+\mathrm{C}\tan^2\beta_2) \quad . \quad . \quad . \quad . \quad . \quad (31)
$$

The value of C, when written out fully, is given by

$$C = \tfrac{3}{4}(n_2-n_0)^{-2}\{-(n_2^2-n_0n_2+n_0^2)+(n_0^2+n_2^2)\,\mathbf{M}_1+(n_0^2-3n_0n_2+n_2^2)\,\mathbf{M}_1^2\}. \quad (32)$$

§ 6. *The Convergency Factor and the Singular Inclination for a Single Refracting Surface.*

Having now obtained expressions (24 and 26) for the longitudinal spherical aberration, which are correct to the second order of aberrations, when expansion in powers of $\sin\alpha_0$ or $\tan\alpha_0$ is legitimate and rapidly convergent, we have now to enquire how far the same expression remains valid as $\mathbf{M}$ increases, in which case we know that '**B** or **B** increases without limit and the convergency fails, even for comparatively small values of α_0.

Here it will be convenient to introduce two definitions :—

I. We shall call *singular inclination* the value λ (see § 3) of α_0 for which the emergent ray is parallel to the axis.

II. The factor $1-\sin^2\alpha_0/\sin^2\lambda$ (or $1-\tan^2\alpha_0/\tan^2\lambda$ if we are dealing with tangents) we shall call the *convergency factor*. If we multiply Δx by the convergency factor we remove those singularities of Δx which are instrumental in causing critical failure of convergency.

To find the singular inclination and convergency factor for a single refracting surface, we have to find when $\alpha_2 = 0$.

Going back to the fundamental equations (8) to (11) we have $\alpha_2 = 0$ when $\alpha_0 = \lambda$, where

$$\lambda = \psi_2-\psi_0$$

which leads to

$$\begin{aligned}\sin\lambda &= \sin\psi_2\cos\psi_0-\sin\psi_0\cos\psi_2\\ &= (n_0x_0\sin\lambda/n_2r_1)\sqrt{\{1-(x_0\sin\lambda/r_1)^2\}}-(x_0\sin\lambda/r_1)\sqrt{\{1-(n_0x_0\sin\lambda/n_2r_1)^2\}}.\end{aligned}$$

Hence, either $\sin\lambda = 0$, which obviously refers to the axial ray, a trivial and (for our purpose) irrelevant solution, or

$$r_1/x_0 = (n_0/n_2)\sqrt{\{1-(x_0\sin\lambda/r_1)^2\}}-\sqrt{\{1-(n_0x_0\sin\lambda/n_2r_1)^2\}}. \quad (33)$$

On rationalising (33) leads to

$$\begin{aligned}4n_0^2\sin^2\lambda/n_2^2 &= 4r_1^2/x_0^2-(1+r_1^2/x_0^2-n_0^2/n_2^2)^2\\ &= -(1+r_1/x_0-n_0/n_2)(1+r_1/x_0+n_0/n_2)(1-r_1/x_0+n_0/n_2)(1-r_1/x_0-n_0/n_2). \quad (34)\end{aligned}$$

This gives the singular inclination.

If we write

$$R \equiv (1+r_1/x_0-n_0/n_2)(1+r_1/x_0+n_0/n_2)(1-r_1/x_0+n_0/n_2)(1-r_1/x_0-n_0/n_2), \quad (35)$$

it follows that

$$1+4n_0^2\sin^2\alpha_0/n_2^2\mathrm{R} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (36)$$

is the required convergency factor.

If the formulæ of § 5 are to get accurately over the failure of convergency, this convergency factor should be identical with

$$1+\mathbf{B}\sin^2\alpha_0/\mathbf{M}_1^2,$$

that is, we should have

$$\mathbf{B} = 4n_0^2\mathbf{M}_1^2/n_2^2\mathrm{R}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

which, when written out, becomes

$$\mathbf{B} = 4n_0^2n_2^2\mathbf{M}_1^2(1-\mathbf{M}_1)^4/(n_2-n_0)^2(1-2\mathbf{M}_1)(n_2+n_0-2n_0\mathbf{M})(n_2+n_0-2n_2\mathbf{M}_1). \quad (38)$$

This does not agree with the previously found value for **B**, being of fractional form in $\mathbf{M}_1$. It does lead to **B** becoming infinite of the order $\mathbf{M}_1^3$ when $\mathbf{M}_1$ tends to infinity, but it indicates an infinity of **B** (and therefore a critical failure of convergency) at three other places, namely when $\mathbf{M}_1 = \frac{1}{2}$, $\frac{1}{2}(n_0+n_2)/n_0$, $\frac{1}{2}(n_0+n_2)/n_2$, at none of which does a failure of convergency really occur, as can readily be verified.

The reason for this is made clearer by geometrical reasoning as follows :—

Let $\mathrm{P}_1\mathrm{Q}_1$ (fig. 3) be a ray which is parallel to the axis in medium 2. To make the figure easier and the quantities dealt with positive, the refraction has been taken from a denser to a rarer medium, so that $n_0/n_2 > 1$.

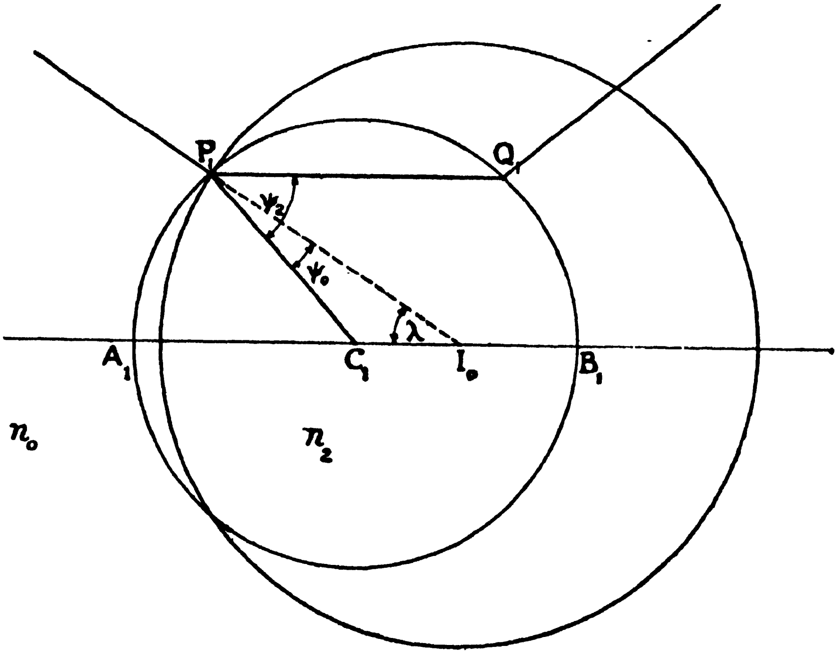


Fig. 3.

In the triangle $\mathrm{I}_0\mathrm{C}_1\mathrm{P}_1$ of fig. 3 we have $\sin\psi_2/\sin\psi_0 = \mathrm{I}_0\mathrm{P}_1/x_0$. But

$$\sin\psi_2/\sin\psi_0 = n_0/n_2\,; \quad \text{hence} \quad \mathrm{I}_0\mathrm{P}_1 = n_0x_0/n_2.$$

The point P_1 and singular inclination λ can therefore be constructed geometrically as follows.

With I_0 as centre and radius $n_0.C_1I_0/n_2$ describe a circle meeting the refracting surface at P_1; $C_1I_0P_1$ is the angle λ required.

This angle λ approaches zero, that is, we get a critical failure of convergency, when the two circles approach contact at A_1. The limiting case is, therefore, when $A_1I_0 = n_0 . C_1I_0/n_2$ or I_0 divides A_1C_1 *externally* in the ratio $n_0 : n_2$.

When this happens $r_1+x_0 = n_0x_0/n_2$ leading to $\mathbf{M}_1 = \infty$, a case of true critical failure. But clearly, by symmetry, we get a precisely similar result when P_1Q_1 is due to a ray entering the surface at Q_1, and travelling backward through medium 2. In this case the limiting position of I_0 divides B_1C_1 externally in the ratio $n_0 : n_2$ and is defined by $r_1-x_0 = -n_0x_0/n_2$ or $\mathbf{M}_1 = \frac{1}{2}$.

This case would correspond, analytically, to $\alpha_2 = \pi$, and the corresponding equation (33) would become

$$-r_1/x_0 = (n_0/n_2)\sqrt{\{1-(x_0 \sin\lambda/r_1)^2\}} - \sqrt{\{1-(n_0x_0 \sin\lambda/n_2r_1)^2\}}.$$

Now if we examine (34) we find that in the process of clearing roots, r_1/x_0 appears squared in the final result, which accordingly includes both $\alpha_2 = 0$ and $\alpha_2 = \pi$. If we write $r_1^2/x_0^2 = u$, then we should really write equation (33) in the form

$$\sqrt{u} = (n_0/n_2)\sqrt{(1-\sin^2\lambda/u)} - \sqrt{(1-n_0^2\sin^2\lambda/n_2^2u)}, \quad . \quad . \quad . \quad . \quad (39)$$

and the two cases are discriminated by assigning to $\sqrt{u}$ one or the other sign. One of these cases is necessarily irrelevant since refraction at the posterior surface of the sphere is physically excluded.

Further, if we consider the other two values which make $R = 0$, viz.,

$$\mathbf{M}_1 = (n_0+n_2)/2n_0 \text{ and } \mathbf{M}_1 = (n_0+n_2)/2n_2,$$

they correspond to

$$r_1+x_0 = -n_0x_0/n_2$$

and

$$r_1-x_0 = n_0x_0/n_2,$$

i.e., to positions of I_0 in which it divides A_1C_1 and B_1C_1 *internally* in the ratio $n_0 : n_2$. But these belong geometrically to the limit of cases in which the incident and refracted rays lie on opposite sides of the normal, *i.e.*, to a negative refractive index. And indeed they are obtained from the two previous points by reversing the sign of n_0/n_2.

Here again, examination of (34) shows that n_0/n_2 appears squared in it. Therefore (34) includes the cases in question. These, however, may be obtained analytically by changing ψ_0, or ψ_2, into its supplement, *i.e.*, by reversing the sign of $\cos\psi_0$ or $\cos\psi_2$ or by changing the determination of the sign of one or other of the square roots on the right-hand side of (39).

The cases are discriminated by the vanishing of these square roots, which occurs when $\cos\psi_0$ or $\cos\psi_2 = 0$.

It appears, therefore, that the vanishing of the factors in the denominator of (38) is wholly irrelevant, and, if we adopted for **B** the value given on the right-hand side of that equation, we should thereby be introducing, in the neighbourhood of $\mathbf{M}_1 = \frac{1}{2}$, $(n_0+n_2)/2n_0$, $(n_0+n_2)/2n_2$ entirely irrelevant singularities, which would make the formula worthless.

The question arises, what is the range of values of $\mathbf{M}_1$ for which the equation

$$4n_0^2 \sin^2\lambda/n_2^2 = -\mathrm{R}$$

is valid and legitimate?

If we start from $\mathbf{M}_1 = \infty$, which corresponds to a real case, the signs of the square roots in (39) are well determined, and the correspondence between $\sin^2\lambda$ and $\mathbf{M}_1$ is unique and definite and can be continued until we reach a point where one case passes into another. These cases we have found to be the branch-points of the three square roots, namely:—

$$u = 0, \qquad \cos\psi_0 = 0 \quad \text{and} \quad \cos\psi_2 = 0.$$

$$u = 0 \quad \text{leads to} \quad x_0 = \infty \text{ or } \mathbf{M}_1 = 0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad \text{(A)}$$

$\cos\psi_0 = 0$ leads to $\sin^2\lambda = u$, or, using the first form of (34)

$$4n_0^2 u/n_2^2 = 4u-(1+u-n_0^2/n_2^2)^2,$$

i.e., $(1-u-n_0^2/n_2^2)^2 = 0$, that is $u = 1-n_0^2/n_2^2$, leading to

$$x_0 = \pm r_1\,(1-n_0^2/n_2^2)^{-\frac{1}{2}} \text{ and } \mathbf{M}_1 = \{1 \pm (n_2-n_0)/\sqrt{(n_2^2-n_0^2)}\}^{-1} \quad . \quad . \quad . \quad \text{(B)}$$

$\cos\psi_2 = 0$ leads to $\sin^2\lambda = n_2^2 u/n_0^2$, that is, to

$$u = n_0^2/n_2^2-1,\ x_0 = \pm r_1\,(n_0^2/n_2^2-1)^{-\frac{1}{2}},\ \mathbf{M}_1 = \{1 \pm (n_2-n_0)/\sqrt{(n_0^2-n_2^2)}\}^{-1} \quad . \quad . \quad \text{(C)}$$

If $n_0 > n_2$, both values of $\mathbf{M}_1$ given by (B) are imaginary. The values given by (C) are both positive, $\mathbf{M}_1 = \{1+(n_2-n_0)/\sqrt{(n_0^2-n_2^2)}\}^{-1}$ being the greater.

The range over which we can travel without ambiguity is, therefore, from $\mathbf{M}_1 = +\infty$ to $\mathbf{M}_1 = \{1+(n_2-n_0)/\sqrt{(n_0^2-n_2^2)}\}^{-1}$ and from $\mathbf{M}_1 = -\infty$ to $\mathbf{M}_1 = 0$.

If $n_0 < n_2$, the values of $\mathbf{M}_1$ given by (C) are imaginary, those given by (B) are positive, and $\mathbf{M}_1 = \{1-(n_2-n_0)/\sqrt{(n_2^2-n_0^2)}\}^{-1}$ is the greater, so that the range of validity is from $\mathbf{M}_1 = +\infty$ to $\mathbf{M}_1 = \{1-(n_2-n_0)/\sqrt{(n_2^2-n_0^2)}\}^{-1}$ and from $\mathbf{M}_1 = -\infty$ to $\mathbf{M}_1 = 0$.

Within this range $(1+4n_0^2\sin^2\alpha_0/n_2^2\mathrm{R})$ is the correct convergency factor; outside this range it is irrelevant.

It is clear, then, that we cannot find a single formula for the convergency factor, which will hold for all values of the magnification.

Further, if the factor $(1+4n_0^2 \sin^2 \alpha_0/n_2^2\mathrm{R})$ is introduced into the denominator of Δx_2, we no longer obtain expressions of the simple type (24) and (26), and endless complications are introduced when we come to consider a compound system.

Can we make our expression $\mathbf{B}$ given by (23) give a tolerable approximation to $(4n_0^2\mathbf{M}_1^2/n_2^2\mathrm{R})$ for those regions where the denominator factor is really needed, namely for $\mathbf{M}_1$ large, positively or negatively ?

To get the answer to this question we develop $(4n_0^2\mathbf{M}_1^2/n_2^2\mathrm{R})$ in descending powers of $\mathbf{M}_1$.

This is found to be (the most rapid method is to break up first into partial fractions)

$$\tfrac{1}{4}(n_2-n_0)^{-2}\left\{\begin{array}{l} -2n_0n_2\mathbf{M}_1^3-(n_0^2-5n_0n_2+n_2^2)\mathbf{M}_1^2 \\ -[(n_2-n_0)^4+n_0^2n_2^2]\mathbf{M}_1/2n_0n_2-[(n_2-n_0)^3(n_2^3-n_0^3)+n_2^3n_0^3]/4n_0^2n_2^2 \\ \quad +\text{terms in } 1/\mathbf{M}_1, \text{ \&c.} \end{array}\right\}. \quad (40)$$

If we now compare (40) with (23) we find that the most important terms when $\mathbf{M}_1$ is large, namely those in $\mathbf{M}_1^3$ and $\mathbf{M}_1^2$ agree in the two expressions.

We may, therefore, take it that the approximations (21) and (24) which we have seen hold good to the second order when expansion in series is convergent, will probably not be numerically very far out when $\mathbf{M}_1$ has a large value, in which case the normal method of development cannot be used.

It is important, at this stage, and to justify the above assertion, to consider a few numerical examples.

Tables I. and II. give the values of Δx_2 and $\sin \alpha_2$ for a single refracting surface, calculated for a number of values of $\mathbf{M}_1$ and two inclinations in each case. The inclinations are fixed from the perpendicular distance ϖ of A_1 from the incident ray. This, for moderate inclinations, is sensibly the same as the intercept made by the incident ray on the principal plane. ϖ has been given the two values 0·5 and 0·25 in every case, except for $\mathbf{M}_1 = 2$ where $\varpi = 0·5$ leads to a physically impossible value. In this case $\varpi = 0·25$ and $\varpi = 0·125$ have been used to define the ray.

In each case four values have been computed (1) the correct one, from trigonometrical calculation ; (2) the values given by formulæ (21) and (24)—these are shown in the column headed "fractional formula" ; (3) the values obtained by expansion in series, up to the optician's first order of aberrations inclusive, that is including $\sin^3 \alpha_0$ in the development of Δx_2 and $\sin \alpha_2$—these are shown in the column headed "first order" ; (4) the same series carried to the second order of aberrations inclusive, *i.e.*, to the terms involving $\sin^5 \alpha_0$—these are shown in the column headed "second order." It should be noted that these first and second order approximations are the most accurate that can be obtained, much more so than more usual ones, proceeding in powers of $\sin \alpha_2$ or $\tan \alpha_2$.

TABLE I.—Values of Δx_2 for Single Refracting Surface.

M.	ϖ.	First order.	Second order.	Fractional formula.	True.
10	0·5	− 33·35293	− 107·6252	+ 27·18563	+ 26·66794
10	0·25	− 8·33823	− 12·98025	− 18·81008	− 18·85273
2	0·25	− 0·250000	− 0·378906	− 0·516129	− 0·527526
2	0·125	− 0·062500	− 0·070557	− 0·071749	− 0·071781
0·5	0·5	− 0·015625	− 0·016541	− 0·016598	− 0·016611
0·5	0·25	− 0·003906	− 0·003963	− 0·003964	− 0·003963
0	0·5	− 0·166667	− 0·174769	− 0·175183	− 0·175809
0	0·25	− 0·041667	− 0·042173	− 0·042179	− 0·042189
− 1	0·5	− 1·000000	− 0·947500	− 0·950119	− 0·956680
− 1	0·25	− 0·250000	− 0·246719	− 0·246761	− 0·246838

TABLE II.—Values of Sin α_2 for Single Refracting Surface.

M.	ϖ.	First order.	Second order.	Fractional formula.	True.
10	0·5	+ 0·0250865	+ 0·0454644	+ 0·0576348	+ 0·0610770
10	0·25	− 0·0078936	− 0·0072568	− 0·0071911	− 0·0071828
2	0·25	− 0·2187500	− 0·2065430	− 0·1987180	− 0·1978219
2	0·125	− 0·1210938	− 0·1207123	− 0·1206710	− 0·1206691
0·5	0·5	0·2539063	0·2541962	0·2542195	0·2542230
0·5	0·25	0·1254883	0·1254974	0·1254975	0·1254974
0	0·5	0·1805556	0·1817130	0·1826667	0·1827294
0	0·25	0·0850694	0·0851267	0·0851286	0·0851291
− 1	0·5	0·1250000	0·1299375	0·1311526	0·1314354
− 1	0·25	0·0531250	0·0532793	0·0532873	0·0532884

It appears from the above that the fractional formulæ are not merely equal, but appreciably superior to the second order formulæ, and this not merely in cases such as those of the three first entries in Table I., in which the convergency of the series for Δx_2 is either absent or slow, but in *every case* where the fractional or second order formulæ differ sensibly from the true value. (Clearly a divergence of 1 in the last place cannot be claimed as significant, for the last figure in Tables I. and II. is probably not correct within ± 2, in some cases.) An estimate of the range of the formula can be obtained from the fact that in the cases, $\varpi = 0·5$, $M_1 = 10$ and 2, the angles of incidence were 52° 34′ and 48° 35′ respectively, and, for the other values, angles of incidence of 20° and 30° are quite common.

In view of this the accuracy of the results is surprising and, from the point of view of the further applications of the method, most encouraging.

§ 7. *Combination of Two Systems.*

Call the systems 1 and 3, and the initial, intermediate and final media 0, 2, 4.

f_1, f_3 are the focal lengths of the systems, as defined in § 2. M_1, $\mathbf{M}_1$ are the transverse and ray magnifications in the first system, M_3, $\mathbf{M}_3$ in the second system.

$M_3+\Delta M_3$, $\mathbf{M}_3+\Delta\mathbf{M}_3$ refer to the transverse and ray magnifications in the second system when I_2, the true intersection of ray 2 with the axis, is taken as the object point for the second refraction (instead of J_2, which refers to transverse and ray magnifications M_3, $\mathbf{M}_3$).

Using the notation of §§ 2, 5, we assume

$$\Delta x_2 = n_2 f_1\,(A_1 t_2^2+E_1 t_2^4)/(1+B_1 t_2^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (41)$$

$$q_2 = t_2\,(1+B_1 t_2^2)/(1+C_1 t_2^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (42)$$

where $q_2 = \genfrac{}{}{0pt}{}{\sin}{\tan}\{\alpha_2\}$, $t_2 = \left\{\begin{matrix}\sin\gamma_2\\ \tan\beta_2\end{matrix}\right\}$, and B_1, C_1 have suitable forms according as sines or tangents are considered. The constant E_1 is zero if the systems reduce to single refracting surfaces. Its form in the more general case will be discussed later.

If we denote by $\Delta_1 x_4$ that part of Δx_4 which is due to Δx_2 and by $\Delta_3 x_4$ that which is introduced by the aberrations proper to the system 3,

$$\Delta_1 x_4 = -n_4 f_3\,\Delta M_3$$

where ΔM_3 is obtained from Δx_2 by means of

$$\Delta x_2 = n_2 f_3\,\{1/(M_3+\Delta M_3)-1/M_3\},$$

leading to

$$\Delta M_3 = -M_3^2\,\Delta x_2/(n_2 f_3+M_3\,\Delta x_2).$$

Thus

$$\Delta_1 x_4 = n_4 f_1 M_3^2\,(A_1 t_2^2+E_1 t_2^4)/\{1+t_2^2\,(B_1+M_3 A_1 f_1/f_3)\} . \quad . \quad . \quad . \quad . \qquad (43)$$

Again

$$\begin{aligned}\Delta_3 x_4 &= \frac{n_4 f_3\,\{(A_3+\Delta A_3)\,q_2^2\,(\mathbf{M}_3+\Delta\mathbf{M}_3)^{-2}+(E_3+\Delta E_3)\,q_2^4\,(\mathbf{M}_3+\Delta\mathbf{M}_3)^{-4}\}}{1+(B_3+\Delta B_3)\,q_2^2\,(\mathbf{M}_3+\Delta\mathbf{M}_3)^{-2}}\\ &= n_4 f_3\,[\{A_3\mathbf{M}_3^{-2}+\Delta M_3\,d\,(A_3\mathbf{M}_3^{-2})/dM_3\}q_2^2+E_3 t_2^4\mathbf{M}_3^{-4}]/(1+B_3 t_2^2\mathbf{M}_3^{-2}), \qquad (44)\end{aligned}$$

retaining only terms of second order in t_2^2. Writing now for q_2^2 in the above its "first order" equivalent $t_2^2+2\,(B_1-C_1)\,t_2^4$, we have

$$\begin{aligned}\Delta x_4/n_4 &= \Delta_1 x_4/n_4+\Delta_3 x_4/n_4\\ &= \frac{\begin{array}{l}\{f_1 M_3^2 A_1+f_3 A_3/\mathbf{M}_3^2\}\,t_2^2+t_2^4\,[f_1\{B_3 A_1 M_3^2\mathbf{M}_3^{-2}+E_1 M_3^2-A_1 M_3^2\,d\,(A_3\mathbf{M}_3^{-2})/dM_3\\ \quad +A_1 A_3 M_3\mathbf{M}_3^{-2}\}+f_3\{B_1 A_3\mathbf{M}_3^{-2}+2A_3\,(B_1-C_1)\,\mathbf{M}_3^{-2}+E_3\mathbf{M}_3^{-4}\}]\end{array}}{\{1+t_2^2\,(B_1+A_1 M_3\,f_1/f_3)\}\,\{1+B_3 t_2^2/\mathbf{M}_3^2\}} . \quad . \quad . \quad . \quad . \qquad (45)\end{aligned}$$

Remembering that $t_2 = \mathbf{M}_3 t_4$ and retaining only the first two terms of the denominator product

$$\Delta x_4/n_4 = (f_{13}A_{13}t_4^2 + f_{13}E_{13}t_4^4)/(1 + B_{13}t_4^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad (46)$$

where

$$f_{13}A_{13} = f_3A_3 + f_1A_1M_3^2\mathbf{M}_3^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (47)$$

$$B_{13} = B_3 + B_1\mathbf{M}_3^2 + A_1M_3\mathbf{M}_3^2 f_1/f_3 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (48)$$

$$\begin{aligned} f_{13}E_{13} &= f_3E_3 + f_1E_1M_3^2\mathbf{M}_3^4 \\ &+ f_1A_1\mathbf{M}_3^2\{B_3M_3^2 + A_3M_3 - M_3^2\mathbf{M}_3^2\,d\,(A_3\mathbf{M}_3^{-2})/dM_3\} \\ &+ f_3A_3\mathbf{M}_3^2\{3B_1' - 2C_1\}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (49) \end{aligned}$$

Again

$$q_4 = \frac{q_2(\mathbf{M}_3 + \Delta\mathbf{M}_3)^{-1}\{1 + (B_3 + \Delta B_3)\,q_2^2(\mathbf{M}_3 + \Delta\mathbf{M}_3)^{-2}\}}{1 + (C_3 + \Delta C_3)\,q_2^2(\mathbf{M}_3 + \Delta\mathbf{M}_3)^{-2}},$$

and retaining only terms of order t_2^3

$$\begin{aligned} q_4 &= q_2(\mathbf{M}_3 + \Delta\mathbf{M}_3)^{-1}(1 + B_3t_2^2\mathbf{M}_3^{-2})/(1 + C_3t_2^2\mathbf{M}_3^{-2}) \\ &= t_2\mathbf{M}_3^{-1}(1 - \Delta\mathbf{M}_3/\mathbf{M}_3)(1 + B_1t_2^2)(1 + B_3t_2^2\mathbf{M}_3^{-2})/\{(1 + C_1t_2^2)(1 + C_3t_2^2\mathbf{M}_3^{-2})\} \\ &= t_2\mathbf{M}_3^{-1}(1 + B_1t_2^2 + B_3t_2^2\mathbf{M}_3^{-2} + t_2^2A_1M_3f_1/f_3)/(1 + C_1t_2^2 + C_3t_2^2\mathbf{M}_3^{-2}) \;. \\ &= t_4(1 + B_{13}t_4^2)/(1 + C_{13}t_4^2), \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (50) \end{aligned}$$

where B_{13} has the value given by (48) and

$$C_{13} = C_3 + C_1\mathbf{M}_3^2. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (51)$$

The equations (47), (48), (49), (51), give the constants for the combined system in terms of those for the components. It may appear at first sight as if the choice of the constants B_{13} and E_{13} had been arbitrary, for clearly, if λ be any quantity,

$$\{f_{13}A_{13}t_4^2 + f_{13}(E_{13} + \lambda A_{13})\,t_4^4\}/\{1 + (B_{13} + \lambda)\,t_4^2\}$$

will give a development equally valid to the second order. But, if we do this, and we wish to preserve the simple character of the relation (48) giving the B for the combination, λ will have to be a linear function of A_1, A_3, B_1, B_3. λA_{13} must then necessarily contain terms of one or other of the forms A_1B_1, A_1^2, A_3B_3, A_3^2. Thus the new E will contain such terms and will no longer be of type (49) which is linear in the aberration coefficients of each system taken separately. Thus the lineo-linear type of equation for E_{13} requires $\lambda = 0$.

We note also that the equations of combination are identical in form, whether we are dealing with the sine or the tangent of the inclination as argument.

§ 8. *Nature of the Quantities* A, B, C, E *in the General Case of any System.*

In the case of the single refracting surface we found that A, B, C were polynomials of degrees 4, 3, 2 in $\mathbf{M}$, and that E was identically zero.

In addition, for such a surface, equations (23), (25) and (27) show that the coefficient of $\mathbf{M}_1^4$ in A is n_0/n_2 times the coefficient of $\mathbf{M}_1^3$ in $\mathbf{B}$ or B. This may be otherwise stated in the form :—

$$A_1 - n_0\mathbf{M}_1B_1/n_2, \text{ i.e., } A_1 - M_1B_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(I)}$$

is of the fourth degree only in appearance and reduces to an expression of the third degree in $\mathbf{M}_1$ or M_1.

The same holds good for $A_1 - M_1\mathbf{B}_1$, so this result is independent of whether the sine or tangent is taken as argument. The same remark applies to all the results of the present section and to the other invariant relations shortly to be proved. We may therefore conveniently state it here once for all.

If we now refer to the equations (47), (48), and remember that in any combination—

$$M_{13}/f_3 = M_3/f_{13} - 1/f_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (52)$$

and

$$1/f_1M_{13} = 1/f_{13}M_1 - 1/f_3 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (53)$$

with the corresponding equations

$$n_0\mathbf{M}_{13}/f_3 = n_2\mathbf{M}_3/f_{13} - n_4/f_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (54)$$

$$n_4/f_1\,\mathbf{M}_{13} = n_2/f_{13}\,\mathbf{M}_1 - n_0/f_3 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (55)$$

and the obvious conditions

$$\mathbf{M}_{13} = \mathbf{M}_1\mathbf{M}_3\,; \qquad M_{13} = M_1M_3,$$

we note first that, if A_3 is a quartic in $\mathbf{M}_3$ it is also a quartic in M_{13} or $\mathbf{M}_{13}$, and that if A_1 is a quartic in $\mathbf{M}_1$, it becomes, on multiplication by $M_3^2\mathbf{M}_3^2$, a quartic in M_{13} or $\mathbf{M}_{13}$, since $M_3 = n_2\mathbf{M}_3/n_4$, and therefore $\mathbf{M}_1^rM_3^2\mathbf{M}_3^2 = n_2^2\mathbf{M}_{13}^r\mathbf{M}_3^{4-r}/n_4^2$, which makes every term in $A_1M_3^2\mathbf{M}_3^2$ a quartic in $\mathbf{M}_{13}$, because $\mathbf{M}_3$ is a linear function of $\mathbf{M}_{13}$ and $4-r$ is here zero or positive.

(47) then shows that A_{13} will be a quartic function of $\mathbf{M}_{13}$, if A_1 and A_3 are quartic functions of $\mathbf{M}_1$ and $\mathbf{M}_3$ respectively. But we know this to be the case for a single refracting surface. Hence it holds good of any system compounded of such surfaces.

Now consider (48). If B_3 is a cubic in $\mathbf{M}_3$ it becomes a cubic in $\mathbf{M}_{13}$.

Again, if $A_1 - M_1B_1 =$ a cubic U_1 in $\mathbf{M}_1$

$$\begin{aligned}\mathbf{M}_3^2(B_1 + A_1M_3f_1/f_3) &= \mathbf{M}_3^2(B_1\{1 + M_{13}f_1/f_3\} + U_1M_3f_1/f_3)\\ &= \mathbf{M}_3^2M_3f_1\,(B_1/f_{13} + U_1/f_3),\end{aligned}$$

using (52), and by reasoning similar to the one given for A_{13} the last expression is a cubic in $\mathbf{M}_{13}$.

Hence B_{13} is a cubic in $\mathbf{M}_{13}$.

Consider now $A_{13}-M_{13}B_{13}$. This is found, after some reductions, and using (52), to be

$$(f_3/f_{13})(A_3-M_3B_3)+f_3B_3/f_1+M_3\mathbf{M}_3^2(A_1-M_1B_1).$$

Of the above terms, $A_3-M_3B_3$ is a cubic in $\mathbf{M}_3$ and therefore also a cubic in $\mathbf{M}_{13}$. B_3 is a cubic in $\mathbf{M}_{13}$. $A_1-M_1B_1$ is a cubic in $\mathbf{M}_1$ and when multiplied by $M_3\mathbf{M}_3^2$ becomes a cubic in $\mathbf{M}_{13}$.

Hence if the condition (I) holds good for the components, it also holds good for the resultant system. But we have seen that it holds for a single refracting system; thus it holds for any combination. Also B will be a cubic in M for any system.

As regards C, examination of (51), remembering that for a single surface C_3 and C_1 are quadratics in $\mathbf{M}_3$, $\mathbf{M}_1$ respectively, leads immediately to the conclusion that C is a quadratic in $\mathbf{M}$ for any system.

We now come to the coefficient E. Here the single refracting surface gives no precedent for E_1 and E_3. Let us examine the other terms in E_{13}. These can be written in the form $f_1A_1\mathbf{M}_3^2M_3^2(B_3-dA_3/dM_3)+3A_3\mathbf{M}_3^2(f_3B_1+f_1M_3A_1)-2f_3A_3\mathbf{M}_3^2C_1$, and, using $A_1-B_1M_1=U_1$; $A_3-B_3M_3=U_3$, where U_1, U_3 are then cubics in M_1, M_3 respectively, this is found to reduce to

$$-f_1A_1\mathbf{M}_3^2M_3^2(M_3\,dB_3/dM_3+dU_3/dM_3)+3A_3\mathbf{M}_3^2M_3f_1(U_1+f_3B_1/f_{13})-2f_3A_3\mathbf{M}_3^2C_1. \quad (56)$$

Now

$$A_1\mathbf{M}_3^2M_3^2 = \text{quartic in } M_{13}.$$

$$M_3\,dB_3/dM_3+dU_3/dM_3 = \text{cubic in } M_3 = \text{cubic in } M_{13}.$$

$$\mathbf{M}_3^2M_3(U_1+f_3\mathbf{B}_1/f_{13}) = \text{cubic in } M_3 = \text{cubic in } M_{13}.$$

$$\mathbf{A}_3 = \text{quartic in } M_{13}.$$

$$\mathbf{M}_3^2C_1 = \text{quadratic in } M_{13}.$$

Hence the three terms in (56) are of form

$$(\text{quartic})(\text{cubic})+(\text{quartic})(\text{cubic})+(\text{quartic})(\text{quadratic}),$$

and this leads to a rational integral polynomial of degree 7 in M_{13}.

Further consideration, however, shows that it is of degree 7 only in appearance, for the terms which can lead to expressions of degree 7 in M_{13} are clearly

$$-f_1\mathbf{M}_3^2M_3^3A_1\,dB_3/dM_3+3\mathbf{A}_3f_1\mathbf{M}_3^2M_3(U_1+f_3B_1/f_{13}),$$

or, dropping the factor f_1

$$\begin{aligned}\mathbf{M}_3{}^2\mathrm{M}_3\{3\mathrm{A}_3(\mathrm{U}_1+f_3\mathrm{B}_1/f_{13})-\mathrm{M}_3{}^2\mathrm{A}_1 d\mathrm{B}_3/d\mathrm{M}_3\}\\ =\mathbf{M}_3{}^2\mathrm{M}_3\{3\mathrm{B}_3\mathrm{M}_3(\mathrm{U}_1+f_3\mathrm{B}_1/f_{13})+3\mathrm{U}_3(\mathrm{U}_1+f_3\mathrm{B}_1/f_{13})\\ -\mathrm{M}_3{}^2\mathrm{M}_1\mathrm{B}_1\, d\mathrm{B}_3/d\mathrm{M}_3-\mathrm{M}_3{}^2\mathrm{U}_1\, d\mathrm{B}_3/d\mathrm{M}_3\}.\end{aligned}$$

The term

$$3\mathrm{U}_3\mathbf{M}_3{}^2\mathrm{M}_3(\mathrm{U}_1+f_3\mathrm{B}_1/f_{13})$$

is clearly of the form cubic × cubic and the terms leading to expressions of 7th degree reduce to

$$\mathbf{M}_3{}^2\mathrm{M}_3\{\mathrm{M}_3\mathrm{U}_1(3\mathrm{B}_3-\mathrm{M}_3\, d\mathrm{B}_3/d\mathrm{M}_3)+\mathrm{M}_3\mathrm{B}_1(3\mathrm{B}_3 f_3/f_{13}-\mathrm{M}_{13}\, d\mathrm{B}_3/d\mathrm{M}_3)\},$$

and since $d\mathrm{M}_{13}/d\mathrm{M}_3=f_3/f_{13}$, this can be written

$$\mathbf{M}_3{}^2\mathrm{M}_3\{\mathrm{M}_3\mathrm{U}_1(3\mathrm{B}_3-\mathrm{M}_3\, d\mathrm{B}_3/d\mathrm{M}_3)+(\mathrm{M}_3\mathrm{B}_1 f_3/f_{13})(3\mathrm{B}_3-\mathrm{M}_{13}\, d\mathrm{B}_3/d\mathrm{M}_{13})\}.$$

But since B_3 is a cubic, in either M_3 or M_{13}, $3\mathrm{B}_3-\mathrm{M}_3 d\mathrm{B}_3/d\mathrm{M}_3$ and $3\mathrm{B}_3-\mathrm{M}_{13} d\mathrm{B}_3/d\mathrm{M}_{13}$ are both quadratics in M_3 or M_{13}.

The above expression therefore reduces to M_3 × sum of two quantities each of form : cubic in M_{13} × quadratic in M_{13}, that is, to a sextic in M_{13}.

We see, therefore, that those terms in E_{13} which do not involve E_1 or E_3 are a polynomial of sixth degree in M_{13} or $\mathbf{M}_{13}$.

It follows that for a *lens* E is necessarily a sextic in the magnification.

Suppose now that E_1 and E_3 are both sextics in $\mathbf{M}_1$, $\mathbf{M}_3$ respectively. Then E_3 is a sextic in M_{13} and $\mathrm{E}_1\mathbf{M}_3{}^4\mathrm{M}_3{}^2$ will also be a sextic in M_{13}, that is E_{13} will again be a sextic in $\mathbf{M}_{13}$.

Hence, since any system is built up of combinations of lenses or single refracting surfaces, we find that E is a sextic polynomial in $\mathbf{M}$ for any system.

Examination of particular cases shows that E is not, in general, divisible by A, so that the vanishing of the latter does not usually involve the disappearance of the second order terms.

§9. *Invariant Relations.*

Certain relations exist between the coefficients A, B, C, E which remain the same in form, whatever the number of refracting surfaces. One of these we have already dealt with, namely the fact that

$$\mathrm{A}-\mathrm{MB}$$

reduces to an expression of the third degree, *i.e.*, the coefficients of highest degree in M in A and B are the same.

This we shall refer to as the *first* invariant relation (I.).

A second invariant relation takes the form

$$\mathrm{B}-\mathrm{C}=\tfrac{3}{8}(1-\mathbf{M}^2)+\tfrac{1}{4}\, d\mathrm{A}/d\mathrm{M}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad \text{(II)}$$

when we use $\tan\beta_{2n}$ as argument, and

$$\mathrm{B}-\mathrm{C} = -\tfrac{1}{8}(1-\mathbf{M}^2)+\tfrac{1}{4}\,d\mathrm{A}/d\mathrm{M}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (\mathrm{II}')$$

when we use $\sin\gamma_{2n}$ as argument.

That these relations hold good for the single refracting surface is readily verified from equations (22), (23), (25), (27) and (32). Suppose now that, for systems 1 and 3 separately, the relation

$$\mathrm{B}-\mathrm{C} = \sigma(1-\mathbf{M}^2)+\tfrac{1}{4}\,d\mathrm{A}/d\mathrm{M}$$

holds, where $\sigma = \tfrac{3}{8}$ or $-\tfrac{1}{8}$ according to the nature of the argument, then from (48) and (51)

$$\begin{aligned}\mathrm{B}_{13}-\mathrm{C}_{13} &= \mathrm{B}_3-\mathrm{C}_3+\mathbf{M}_3^2(\mathrm{B}_1-\mathrm{C}_1)+\mathrm{A}_1\mathrm{M}_3\mathbf{M}_3^2 f_1/f_3\\ &= \sigma(1-\mathbf{M}_{13}^2)+\tfrac{1}{4}\,d\mathrm{A}_3/d\mathrm{M}_3+\tfrac{1}{4}\mathbf{M}_3^2(d\mathrm{A}_1/d\mathrm{M}_1+4\mathrm{A}_1\mathrm{M}_3 f_1/f_3), \quad . \quad . \qquad (57)\end{aligned}$$

and from (47)

$$f_{13}\,d\mathrm{A}_{13}/d\mathrm{M}_{13} = f_3\,d\mathrm{A}_3/d\mathrm{M}_{13}+f_1\{\mathrm{M}_3^2\mathbf{M}_3^2\,d\mathrm{A}_1/d\mathrm{M}_{13}+\mathrm{A}_1\,d(\mathrm{M}_3^2\mathbf{M}_3^2)/d\mathrm{M}_{13}\}.$$

But

$$d\mathrm{M}_{13} = (f_3/f_{13})\,d\mathrm{M}_3 = (f_1/f_{13})\,\mathrm{M}_3^2\,d\mathrm{M}_1.$$

Hence

$$d\mathrm{A}_{13}/d\mathrm{M}_{13} = d\mathrm{A}_3/d\mathrm{M}_3+\mathbf{M}_3^2\,d\mathrm{A}_1/d\mathrm{M}_1+(f_1\mathrm{A}_1/f_3)\,d(\mathrm{M}_3^2\mathbf{M}_3^2)/d\mathrm{M}_3,$$

and since $\mathbf{M}_3/\mathrm{M}_3 =$ const., the last differential coefficient is $4\mathrm{M}_3\mathbf{M}_3^2$.

Using this result (57) becomes

$$\mathrm{B}_{13}-\mathrm{C}_{13} = \sigma(1-\mathbf{M}_{13}^2)+\tfrac{1}{4}\,d\mathrm{A}_{13}/d\mathrm{M}_{13},$$

which is of the same form as the equation we started from. Hence, if the two components of the compound system satisfy the second invariant relation, the resultant system also satisfies it. But we have seen that the relation holds good for single refracting surfaces—hence it holds good universally.

It should be noted that the second invariant relation is really a *first order* relation and connects the first order aberration of the inclination of a ray, with the first order longitudinal spherical aberration.

§ 10. *The Constants* A, B, C, E *for an Optical System Reversed and for Negative Lenses.*

Certain important general relations are found to hold between the constants A, B, C, E for rays going through an optical system and the corresponding constants A′, B′, C′, E′, for the same system reversed, and by making use of them we can obtain either set from the other.

We arrive most simply at these relations as follows:—If after traversing the system we retrace our steps, the result is equivalent to compounding the system with itself reversed, with the difference that, in the second set of refractions, the measurement of length parallel to the axis is reversed in direction. An examination of the equations (41) *et seq.*, § 7, on which the formulæ of combination are based, shows that this is analytically equivalent to changes in the sign of the focal length in the second set of refractions.

We have therefore

$$f_1 = f, \quad f_3 = -f, \quad \mathrm{M}_1 = \mathrm{M}, \quad \mathrm{M}_3 = 1/\mathrm{M}, \quad \mathbf{M}_1 = \mathbf{M}, \quad \mathbf{M}_3 = 1/\mathbf{M}, \quad \mathbf{M}_{13} = \mathrm{M}_{13} = 1,$$

and we also find that $f_{13} = \infty$. But $f_{13}\mathrm{A}_{13}$, and $f_{13}\mathrm{E}_{13}$ have definite limiting values, and as f_{13} does not otherwise explicitly enter into the equations of combination, no difficulty arises on that account.

Now, after retracing our steps in this way, we necessarily arrive at a perfect image, so that $\Delta x_4 \equiv 0$ and $\tan \alpha_4 = \tan \beta_4$, leading to

$$f_{13}\mathrm{A}_{13} \equiv 0,$$

$$f_{13}\mathrm{E}_{13} \equiv 0,$$

and

$$\mathrm{B}_{13} - \mathrm{C}_{13} \equiv 0.$$

These lead to the following identical relations

$$\mathrm{A}(\mathrm{M})/\mathrm{M}^2\mathbf{M}^2 - \mathrm{A}'(\mathrm{M}^{-1}) \equiv 0, \quad \ldots \ldots \quad (58)$$

$$\mathrm{E}(\mathrm{M})/\mathrm{M}^2\mathbf{M}^4 - \mathrm{E}'(\mathrm{M}^{-1}) + \mathrm{A}(\mathrm{M})\,\mathbf{M}^{-2}\{\mathrm{B}'(\mathrm{M}^{-1})\,\mathrm{M}^{-2} + 3\mathrm{A}'(\mathrm{M}^{-1})\,\mathrm{M}^{-1} + d\mathrm{A}'(\mathrm{M}^{-1})/d\mathrm{M}\}$$
$$- \mathrm{A}'(\mathrm{M}^{-1})\,\mathbf{M}^{-2}\{3\mathrm{B}(\mathrm{M}) - 2\mathrm{C}(\mathrm{M})\} \equiv 0, \quad \ldots \ldots \quad (59)$$

$$\mathrm{B}'(\mathrm{M}^{-1}) - \mathrm{C}'(\mathrm{M}^{-1}) + \mathbf{M}^{-2}\{\mathrm{B}(\mathrm{M}) - \mathrm{C}(\mathrm{M})\} - \mathrm{A}(\mathrm{M})\,\mathrm{M}^{-1}\mathbf{M}^{-2} \equiv 0. \quad \ldots \quad (60)$$

Equation (58) may be written in either of the two forms

$$\mathrm{A}(\mathrm{M}) \equiv \mathrm{M}^2\mathbf{M}^2\mathrm{A}'(\mathrm{M}^{-1})$$

$$\mathrm{A}(\mathbf{M})/n_0^2 \equiv \mathbf{M}^4\mathrm{A}'(\mathbf{M}^{-1})/n_2^2. \quad \ldots \ldots \quad (\mathrm{III})$$

This we shall refer to as the third invariant relation. It shows that, if we divide A by the square of the initial refractive index, the coefficients of powers of $\mathbf{M}$ equidistant from the beginning and end of the development are interchanged by reversing the system. Equation (59) becomes on multiplying up by $\mathrm{M}^2\mathbf{M}^4$, using (58) and simplifying

$$\mathrm{E} - \mathrm{M}^2\mathbf{M}^4\mathrm{E}' + \mathrm{A}\{\mathbf{M}^2\mathrm{B}' - \mathrm{A}/\mathrm{M} - 3\mathrm{B} + 2\mathrm{C} + d\mathrm{A}/d\mathrm{M}\} \equiv 0, \quad \ldots \quad (61)$$

omitting the arguments M, 1/M of A, B, B′, &c., since no confusion can occur.

Use now the second invariant relation

$$d\mathrm{A}/d\mathrm{M} \equiv 4\mathrm{B} - 4\mathrm{C} - 4\sigma(1 - \mathbf{M}^2),$$

(61) becomes

$$E-M^2\mathbf{M}^4E'+A\{B+\mathbf{M}^2B'-2C-A/M-4\sigma(1-\mathbf{M}^2)\}\equiv 0. \quad . \quad . \quad . \quad (62)$$

Now substitute from (60) for $B+\mathbf{M}^2B'$ the value

$$C+\mathbf{M}^2C'+A/M,$$

and (61) leads to

$$E-M^2\mathbf{M}^4E'+A\{\mathbf{M}^2C'-C-4\sigma(1-\mathbf{M}^2)\}\equiv 0 \quad . \quad . \quad . \quad . \quad . \quad (63)$$

For a single refracting surface, where E, E′ are identically zero, this must lead to

$$\mathbf{M}^2C'-C=4\sigma(1-\mathbf{M}^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{IV})$$

a result which is easily verified from equation (32).

Now consider a system compounded of two systems. For the system direct, we have

$$C_{13}=C_3+\mathbf{M}_3{}^2C_1.$$

Similarly, for the system reversed, change C_3 into C'_1, C_1 into C'_3, $\mathbf{M}_3$ into $1/\mathbf{M}_1$.

$$\begin{aligned} C_{31} &= C'_1+\mathbf{M}_1{}^{-2}C'_3 \\ \mathbf{M}_{13}{}^2C_{31}-C_{13} &= \mathbf{M}_1{}^2\mathbf{M}_3{}^2C'_1-\mathbf{M}_3{}^2C_1+\mathbf{M}_3{}^2C'_3-C_3 \\ &= \mathbf{M}_3{}^2(\mathbf{M}_1{}^2C'_1-C_1)+\mathbf{M}_3{}^2C'_3-C_3 \end{aligned}$$

and using (IV) which we know to be true for a single refracting surface

$$\begin{aligned} \mathbf{M}_{13}{}^2C_{31}-C_{13} &= 4\sigma[\mathbf{M}_3{}^2(1-\mathbf{M}_1{}^2)+1-\mathbf{M}_3{}^2] \\ &= 4\sigma(1-\mathbf{M}_{13}{}^2). \end{aligned}$$

In other words (IV) will hold for the resultant system if it holds for the components, and therefore as in previous similar cases, it holds for any system.

We shall call (IV) the *fourth* invariant relation.

Equation (63) then shows that there exists a *fifth* invariant relation

$$E=M^2\mathbf{M}^4E'. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (\text{V})$$

or

$$E(\mathbf{M})/n_0{}^2=\mathbf{M}^6E'(\mathbf{M}^{-1})/n_2{}^2,$$

so that E possesses a property similar to that of A, previously noticed, viz., if we divide it by the square of the initial refractive index, the coefficients of powers of $\mathbf{M}$ equidistant from the beginning and end of the development in $\mathbf{M}$ are interchanged.

Equation (59) has therefore led us to *two* independent invariant relations.

On the other hand it will be found that (60) leads to no new relation. For if we substitute into it for $B'-C'$ and $B-C$ in virtue of the second invariant relation, and then use the third relation, it becomes an identity.

It gives, however, on using (IV) to eliminate C′,

$$\mathbf{M}^2B'+B = A/M+2C+4\sigma(1-\mathbf{M}^2) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (64)$$

which is a convenient form for calculating B′.

If now A, B, C, E are known for any system, the corresponding quantities are immediately obtainable for the reversed system, A′, B′, C′, E′ being given by equations (III), (64), (IV) and (V) respectively.

This will generally halve the labour of calculations, if it is found desirable to tabulate these constants for a complete set of lenses. It will then be sufficient to start from the equi-convex lens and vary the curvatures in one sense only.

Incidentally we note also that the second invariant relation enables us to find C, so soon as A and B are known, so that only A, B and E require to be calculated.

The aberration constants for a reversed system have a further important application in the case of lenses. Consider a positive lens (fig. 4), the initial ray converging to I_0 and the final ray to I_4. If now we interchange the full and dotted portions of the initial and final rays in fig. 4, we obtain, since here the initial and final media are the same, the case of a ray going through a lens in which the front and back character of the two surfaces have been interchanged. In fact r_1 and r_3 have been interchanged and the sign of the thickness c_2 has been reversed. This leads to a negative lens, of the same *numerical* power as the original positive lens, and with the same mean curvature, but a negative thickness. Such a lens, of course, is not physically realisable, although a part of it can be physically obtained by rotating the wedge *beyond* the intersection V of the two surfaces.

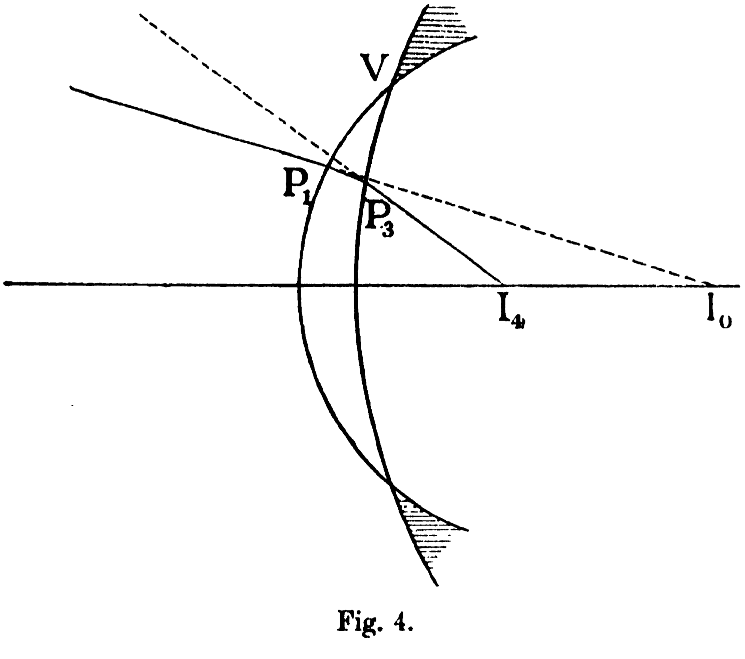


Fig. 4.

But, in the case of the ideally *thin* lenses, where the thickness is zero, the ideally thin positive and negative lenses, having the same mean curvature and numerical power, correspond in this way.

Now I_4 and I_0 are also interchanged. If we consider I_4 as the initial point, then we are really considering a set of rays starting from I_4 in the last medium and travelling backwards through the original positive lens. In other words, the aberration constants for the corresponding negative lens are identical with those for the original positive lens reversed, and the equations (III), (64), (IV) and (V) are applicable to calculate them.

This, again, will greatly diminish the work of calculation. In the case of ideally thin negative lenses, we see that A, B, C, E are directly obtained from the corresponding thin positive lenses. In the case of *thick* negative lenses the corresponding positive lens has a negative thickness.

Now for various reasons it will probably be convenient, in calculating A, B, C, E for lenses, to express them in the form

$$A_c = A_0 + c\,(dA/dc)_0,$$

&c., where c, the thickness, is small, as it usually is in practice, and A_0 refers to an ideally thin lens.

When the formulæ are put in this form, it is perfectly simple to calculate A_{-c}, B_{-c}, &c., and then to obtain the corresponding results for the negative lens with a positive thickness.

§ 11. *Explicit Values of* A, B, C, E *for a Thick Lens* (*Tangent Formula*).

For the purposes of numerical calculation and comparison with correct trigonometrically found values, we have worked out explicitly the form of the expressions A, B, C, E for a thick lens, when we use $\tan\beta_4$ as the argument; the formulæ are expressed in terms of the focal lengths of each surface and of the combination and the thickness does not appear explicitly. The initial and final media being the same $n_0 = n_4$ and we have written $n = n_2/n_0$.

The work of algebraic calculation has been straightforward but extremely heavy, and we therefore omit it here entirely, the object being to publish the results for reference, in case other workers desire to use them for tabulation purposes, but it is hardly to be expected that designing opticians should work direct from the algebraic expressions as they stand.

For the purpose of this section we shall write

$$A = A_0 + A_1M + A_2M^2 + A_3M^3 + A_4M^4.$$

$$B = B_0 + B_1M + B_2M^2 + B_3M^3.$$

$$C = C_0 + C_1M + C_2M^2.$$

$$E = E_0 + E_1M + E_2M^2 + E_3M^3 + E_4M^4 + E_5M^5 + E_6M^6.$$

The suffixes here have a different meaning to that which has been previously ascribed to them, but no confusion is likely to arise on this account.

The values are as follows, f denoting the focal length of the lens, f_1, f_3 the focal lengths of the surfaces, so that

$$f_1 = r_1/(n-1), \qquad f_3 = -r_3/(n-1)$$

$$\mathrm{A}_0 = \frac{1}{2n^2(n-1)^2}\left[-\frac{f_1+f_3}{f_1^4}f^3 + \frac{(n+1)^2 f_3 f^2}{f_1^3} - 2n(1+n+n^2)\frac{f_3 f}{f_1^2} + n^2(n+1)^2\frac{f_3}{f_1} - \frac{n^4 f_3}{f}\right]$$

$$\mathrm{A}_1 = \frac{1}{2n^2(n-1)^2}\left[-\frac{4(f_1+f_3)}{f_1^3 f_3}f^3 + 4(n+1)^2\frac{f^2}{f_1^2} - 4n(1+n+n^2)\frac{f}{f_1} + n^2(n+1)^2\right]$$

$$\mathrm{A}_2 = \frac{1}{2n^2(n-1)^2}\left[-\frac{6(f_1+f_3)}{f_1^2 f_3^2}f^3 + 6(n+1)^2\frac{f^2}{f_1 f_3} + 2n(1+n+n^2)f\left(\frac{1}{f_1}+\frac{1}{f_3}\right)\right]$$

$\mathrm{A}_3 = \mathrm{A}_1$ with f_1 and f_3 interchanged

$\mathrm{A}_4 = \mathrm{A}_0$ " " "

$$\mathrm{B}_0 = \frac{1}{4n^2(n-1)^2}\left[-\frac{2f^3}{f_1^2}\left(\frac{1}{f_1}+\frac{1}{f_3}\right) + 2(n^2-n+1)\frac{f^2}{f_1^2} + n(n-1)^2\frac{f}{f_1} - n^2(n^2-n+1)\right]$$

$$\mathrm{B}_1 = \frac{1}{4n^2(n-1)^2}\left[-\frac{6f^3}{f_1 f_3}\left(\frac{1}{f_1}+\frac{1}{f_3}\right) + 6(n^2+1)\frac{f^2}{f_1 f_3} + n(n-1)^2 f\left(\frac{1}{f_1}+\frac{1}{f_3}\right)\right]$$

$$\mathrm{B}_2 = \frac{1}{4n^2(n-1)^2}\left[-\frac{6f^3}{f_3^2}\left(\frac{1}{f_1}+\frac{1}{f_3}\right) + 6(n^2+n+1)\frac{f^2}{f_3^2} - 3n(n+1)^2\frac{f}{f_3} + 3n^2(n^2-n+1)\right]$$

$\mathrm{B}_3 = \mathrm{A}_4$

$\mathrm{C}_0 = \mathrm{B}_0 - \frac{3}{8} - \frac{1}{4}\mathrm{A}_1$

$\mathrm{C}_1 = \mathrm{B}_1 - \frac{1}{2}\mathrm{A}_2$

$\mathrm{C}_2 = \mathrm{B}_2 + \frac{3}{8} - \frac{3}{4}\mathrm{A}_3$

$$\mathrm{E}_0 = \frac{3}{8(n-1)^4}\left[\begin{aligned} & -\frac{(n^2-n+1)}{n^4}\,\frac{(f_1+f_3)f^5}{f_1^6} + \frac{(n+1)^2}{n^4}f^4\left\{\frac{n}{f_1^4} + \frac{(n^2-n+1)f_3}{f_1^5}\right\} \\ & -\frac{f^3}{n^3}\left\{\frac{n^3+3n^2+n}{f_1^3} + \frac{2(n^4+n^2+1)f_3}{f_1^4}\right\} + \frac{f^2}{n^2}\left\{\frac{2n^3}{f_1^2} + (n+1)^3(n^2-n+1)\frac{f_3}{f_1^3}\right\} \\ & \qquad -f(n^2-n+1)\frac{f_3}{f_1^2} \end{aligned}\right]$$

$$\mathrm{E}_1 = \frac{3}{8(n-1)^4}\left[\begin{aligned} & -\frac{6(n^2-n+1)}{n^4}\,\frac{(f_1+f_3)f^5}{f_3 f_1^5} + \frac{(n+1)^2}{n^4}f^4\left\{\frac{5n}{f_3 f_1^3} + \frac{6(n^2-n+1)}{f_1^4} + \frac{n f_3}{f_1^5}\right\} \\ & -\frac{f^3}{n^3}\left\{\frac{4(n^3+3n^2+n)}{f_3 f_1^2} + \frac{(n+1)^4+8(n^4+n^2+1)}{f_1^3} + (n+1)^4\frac{f_3}{f_1^4}\right\} \\ & +\frac{f^2}{n^2}\left\{\frac{6n^2}{f_3 f_1} + \frac{4(n+1)^2(n^2+1)}{f_1^2} + \frac{2(n-1)^2(n^2+n+1)f_3}{f_1^3}\right\} \\ & -\frac{f}{n}\left\{\frac{2n(2n^2+n+2)}{f_1} + \frac{(n+1)^4 f_3}{f_1^2}\right\} + \frac{n f_3}{f_1}(n+1)^3 \end{aligned}\right]$$

$$
\begin{aligned}
E_2 = \frac{3}{8(n-1)^4}\Bigg[& -\frac{15(n^2-n+1)}{n^4}\frac{f^5(f_1+f_3)}{f_1^4 f_3^2} + \frac{(n+1)^2}{n^4} f^4\left\{\frac{10n}{f_1^2 f_3^2} + \frac{15(n^2-n+1)}{f_1^3 f_3} + \frac{5n}{f_1^4}\right\} \\
& -\frac{f^3}{n^3}\left\{\frac{6(n^3+3n^2+n)}{f_1 f_3^2} + \frac{12(n^4+n^2+1)+4(n+1)^4}{f_1^2 f_3}\right. \\
& \qquad\qquad \left. + \frac{4(n+1)^4+2(n^4+n^2+1)}{f_1^3} + \frac{(n^3+3n^2+n)f_3}{f_1^4}\right\} \\
& +\frac{f^2}{n^2}\left\{\frac{6n^2}{f_3^2} + \frac{6(n+1)^2(n^2+n+1)}{f_1 f_3} + \frac{8(n+1)^2(n^2+n+1)}{f_1^2}\right. \\
& \qquad\qquad \left. + \frac{(n+1)^2(n^2+3n+1)f_3}{f_1^3}\right\} \\
& -\frac{f}{n}\left\{\frac{4n(n+1)^2+n^3-n^2+n}{f_3} + \frac{2(n+1)^4+2(n^2+n+1)(n^2+3n+1)}{f_1}\right. \\
& \qquad\qquad \left. + \frac{2(n^2+n+1)(n^2+3n+1)f_3}{f_1^2}\right\} \\
& + n\{(n+1)^2+4(n^2+n+1)\} + (n+1)^2(n^2+3n+1)\frac{f_3}{f_1} \\
& \qquad\qquad -n^2(n^2+3n+1)\frac{f_3}{f_1}\Bigg]
\end{aligned}
$$

$$
\begin{aligned}
E_3 = \frac{3}{8(n-1)^4}\Bigg[& -\frac{20(n^2-n+1)}{n^4}\frac{f^5(f_1+f_3)}{f_1^3 f_3^3} + \frac{(n+1)^2}{n^4} f^4\left\{\frac{10n}{f_3 f_1}\left(\frac{1}{f_1^2}+\frac{1}{f_3^2}\right)\right. \\
& \qquad\qquad \left. + \frac{20(n^2-n+1)}{f_1^2 f_3^2}\right\} \\
& -\frac{f^3}{n^3}\left\{4(n^3+3n^2+n)\left(\frac{1}{f_1^3}+\frac{1}{f_3^3}\right) + \frac{8(n^4+n^2+1)+6(n+1)^4}{f_1 f_3}\left(\frac{1}{f_1}+\frac{1}{f_3}\right)\right\} \\
& +\frac{f^2}{n^2}\left\{2n^2\left(\frac{f_1}{f_3^3}+\frac{f_3}{f_1^3}\right) + 4(n+1)^4\left(\frac{1}{f_1^2}+\frac{1}{f_3^2}\right) + \frac{12(n+1)^2(n^2+n+1)}{f_1 f_3}\right\} \\
& -\frac{f}{n}\left\{2n(n+1)^2\left(\frac{f_1}{f_3^2}+\frac{f_3}{f_1^2}\right) + [(n+1)^4\right. \\
& \qquad\qquad \left. + 4(n^2+n+1)(n^2+3n+1)]\left(\frac{1}{f_1}+\frac{1}{f_3}\right)\right\} \\
& + 4n(n^2+n+1)\left(\frac{f_1}{f_3}+\frac{f_3}{f_1}\right) + 2(n+1)^2(n^2+3n+1) \\
& - 2n^2(n+1)^2\frac{(f_1+f_3)}{f} + 2n^4\frac{f_1 f_3}{f^2}\Bigg]
\end{aligned}
$$

$E_4 = E_2$ with f_1 and f_3 interchanged.

$E_5 = E_1$,, ,, ,, ,,

$E_6 = E_0$,, ,, ,, ,,

It should be noted carefully that all the above refer to expressions in terms of the tangent of the Gaussian inclination, this being the argument we have used in the numerical work.

§ 12. *Values of* A, B, C, E *for a Thin Lens.*

When the lens is thin, we have the relation

$$\frac{1}{f_1}+\frac{1}{f_3}=\frac{1}{f},$$

which enables the values of § 11 to be considerably simplified.

In this case it is useful to introduce a quantity K such that

$$\mathrm{K}=\frac{\text{mean curvature of the lens}}{\text{power of the lens}}=\frac{f}{2}\left(\frac{1}{r_1}+\frac{1}{r_3}\right).$$

When this is done the constants A, B, C, E take the following forms :—

$$\begin{aligned}\mathrm{A} = -\{(1-\mathrm{M})^2/2n\}\ \{&(n+2)\,[(1-\mathrm{M})\,\mathrm{K}-(1+\mathrm{M})\,(n+1)/(n+2)]^2\\ &+n^3\,(1-\mathrm{M})^2/4\,(n-1)^2-n^2\,(1+\mathrm{M})^2/4\,(n+2)\}.\end{aligned}$$

$$\begin{aligned}\mathrm{B} = {}&(1-\mathrm{M})^2\,\mathrm{K}^2\,\{n-1-(n+2)\,\mathrm{M}\}/2n+(1-\mathrm{M})\,\mathrm{K}\,(1+\mathrm{M}+4\mathrm{M}^2)\,(n+1)/4n\\ &+(1-\mathrm{M})\,\{\mathrm{M}\,(1+\mathrm{M})/4n+(1-\mathrm{M})\,[3n-2n^2-3+\mathrm{M}\,(6n-4n^2-3)]/8\,(n-1)^2\}.\end{aligned}$$

$$\mathrm{C} = -3\,(1-\mathrm{M})^2\,\mathrm{K}^2/2n+3\,(1-\mathrm{M}^2)\,\mathrm{K}\,(n+1)/4n-\tfrac{3}{4}\,(1-\mathrm{M}^2)-\tfrac{3}{8}n\,(1-\mathrm{M})^2/(n-1)^2.$$

$$\mathrm{E} = \frac{3\,(1-\mathrm{M})^2}{128\,(n-1)^4\,n^3}\left[\begin{aligned}&(1+\mathrm{M}^4)\,(-4n^5+8n^4-n^3-4n^2+3n-1)\\ &\quad+\mathrm{M}\,(1+\mathrm{M}^2)\,(8n^6-16n^5+4n^4+4n^3-12n^2+12n-4)\\ &\quad+\mathrm{M}^2\,(-16n^7+32n^6-8n^5-8n^4+10n^3-16n^2+18n-6)\\ &\quad+8\,(n-1)\,(1-\mathrm{M})\,(1+\mathrm{M})\,\mathrm{K}\,\{(1+\mathrm{M})^2\\ &\qquad\qquad(-2n^5+5n^4-2n^3-3n^2+3n-1)\\ &\qquad\qquad+\mathrm{M}\,(2n^6-8n^4+4n^3+2n^2)\}\\ &\quad+8\,(n-1)^2\,(1-\mathrm{M})^2\,\mathrm{K}^2\,\{(1+\mathrm{M})^2\\ &\qquad\qquad(-2n^5+8n^4-7n^3-6n^2+9n-3)\\ &\qquad\qquad-2\mathrm{M}\,n^2\,(n^2-2n-1)\}\\ &\quad+16\,(n-1)^3\,(1+\mathrm{M})\,(1-\mathrm{M})^3\,\mathrm{K}^3\,\{2n^4-4n^3-2n^2+6n-2\}\\ &\quad+16\,(n-1)^4\,(1-\mathrm{M})^4\,\mathrm{K}^4\,(-n^3+3n-1)\end{aligned}\right].$$

We notice that when M = 1 (which gives one of the zeros of A) B, C and E all vanish with it, and also E/A remains finite. Hence in this case, the term $\mathrm{E}t_4{}^4$ will not rise in importance, even when A = 0. But in this case A may have two other real zeros, and these are not zeros of E, so that E plays an important part in the neighbourhood of such zeros.

§ 13. *Numerical Test for a Single Lens.* .

To test the formulæ, a number of longitudinal aberrations were calculated trigonometrically for five positive lenses of refractive index 1·52, unit focal length and thickness $\frac{1}{10}$. The first was a meniscus-shaped lens for which $r_1 = 0·349418$, $r_3 = 1$.

The second was a plano-convex lens, of which the convex side is towards the incoming light. The third was an equi-convex lens. The fourth and fifth were the second and first reversed. If K has the meaning defined in § 12, the values of K for these five lenses are 1·93094, 0·96154, 0, −0·96154 and − 1·93094 respectively, so that they proceed by approximately equal steps of K.

The constants A, B, E were calculated for these five lenses and the longitudinal aberrations computed from the formula. The rays selected met the first principal plane at a distance 0·15 from the axis, corresponding to an aperture *f*/3·4, nearly.

The results are shown in Table III.

TABLE III.—Longitudinal Aberrations of Five Selected Lenses.

M.	Lens 1.			Lens 2.			Lens 3.		
	Formula.	True.	Percentage error.	Formula.	True.	Percentage error.	Formula.	True.	Percentage error.
3	−7·98683	−7·87422	1·41	−1·12450	−1·12814	0·32	−0·32378	−0·32439	0·19
2	−1·55287	−1·52334	1·94	−0·40167	−0·40283	0·29	−0·14240	−0·14282	0·29
0·5	−0·003527	−0·003530	0·09	−0·011697	−0·011705	0·07	−0·033008	−0·033096	0·27
0	−0·06028	−0·06043	0·24	−0·024728	−0·024749	0·08	−0·036194	−0·036213	0·05
−0·5	−0·20902	−0·20946	0·21	−0·083863	−0·083950	0·10	−0·057067	−0·057101	0·06
−1	−0·42692	−0·42772	0·19	−0·18428	−0·18448	0·11	−0·094736	−0·094787	0·05
−2	−1·00383	−1·00531	0·15	−0·48874	−0·48930	0·12	−0·21707	−0·21724	0·08

M.	Lens 4.			Lens 5.		
	Formula.	True.	Percentage error.	Formula.	True.	Percentage error.
3	−0·098969	−0·099086	0·12	−0·10401	−0·10444	0·41
2	−0·046089	−0·046128	0·30	−0·013671	−0·013686	0·11
0·5	−0·073959	−0·074535	0·77	−0·14762	−0·15068	2·03
0	−0·099843	−0·100044	0·20	−0·22172	−0·22324	0·68
−0·5	−0·13967	−0·13984	0·12	−0·32828	−0·32966	0·42
−1	−0·19084	−0·19101	0·09	−0·45753	−0·45887	0·29
−2	−0·32454	−0·32476	0·07	−0·77417	−0·77511	0·12

The mean percentage error of these results is 0·34, so that the formula determines, on the average, the longitudinal aberration correct to 1 part in 300. The bulk of this, however, is contributed by three cases, namely: $M = 3$ and 2 for lens 1 where the aberrations are very large and differ very widely from the usual first and second order approximations, so that, although the error of the formula approaches 2 per cent., it nevertheless represents a great improvement upon these approximations; and $M = 0{\cdot}5$ for lens 5, which corresponds to extreme curvature and highest inclination, so that one of the angles of refraction is as great as 48½ degrees. Even here the table below shows that the formula is an appreciable improvement on the usual second-order approximation. If these three cases are omitted, the mean percentage error works out to be about 0·21, so that in general the formula determines the longitudinal aberration correct to about 1 part in 500.

It is interesting to note what the usual first and second order approximations lead to in a few cases.

	First order approximation.	Percentage error.	Second order approximation.	Percentage error.
Lens 1, $M = 3$ . . .	$-1{\cdot}34012$	$83{\cdot}0$	$-2{\cdot}45432$	$68{\cdot}8$
„ 1, $M = 2$ . . .	$-0{\cdot}48545$	$68{\cdot}1$	$-0{\cdot}82104$	$46{\cdot}1$
„ 1, $M = -2$ . . .	$-1{\cdot}348142$	$34{\cdot}1$	$-0{\cdot}88456$	$12{\cdot}0$
„ 2, $M = 3$ . . .	$-0{\cdot}67356$	$40{\cdot}3$	$-0{\cdot}94436$	$16{\cdot}3$
„ 5, $M = 0{\cdot}5$ . . .	$-0{\cdot}12136$	$19{\cdot}5$	$-0{\cdot}14307$	$5{\cdot}1$

This gives a measure of the numerical improvement effected by the fractional formula whenever the usual method of approximation is seriously out, even though in none of the cases above does the convergency of the series actually fail.

In the above the series are in powers of $\tan\beta_2$. Had they been taken in powers of $\tan\alpha_2$, as is frequently done, the first and second order approximations would have been far worse.

One interesting outcome of these calculations relates to the relative importance of the terms in Et_4^4 and At_4^2. The ratio Et_4^2/A is small in every case taken (of course these exclude the neighbourhood of points where $A = 0$, where naturally E becomes of great importance). But for the set of magnifications taken, the greatest ratio of the second term to the first is less than 0·03 and the mean value of this ratio is only 0·0082, so that, in fact, the E term—although so complicated algebraically—does not exercise any great influence numerically.

This is important, as it shows that, at any rate for lenses, it does not require to be computed with anything like the same order of accuracy which is needed for A and B.

§ 14. *The Singular Inclination and Convergency Factor for any System.*

Referring again to fig. 2 we see that $\lambda = \alpha_2$ and $F_0I_0 = -\Delta x_2$ for rays proceeding through the system reversed and initially parallel.

Thus using accents, as before, to denote the coefficients and inclinations for the system reversed, and noting that the accented coefficients all refer to zero magnification, we have, using *tangents*

$$\tan\lambda = \tan\beta'_2\,(1+\mathrm{B}'_0\tan^2\beta'_2)/(1+\mathrm{C}'_0\tan^2\beta'_2)$$

and

$$-\mathrm{F}_0\mathrm{I}_0 = fn_0\,(\mathrm{A}'_0\tan^2\beta'_2+\mathrm{E}'_0\tan^4\beta'_2)/(1+\mathrm{B}'_0\tan^2\beta'_2).$$

Here the suffixes in the A, B, C, &c., A′, B′, C′, &c., have the same meaning as in § 11.

Thus

$$-1/\mathrm{M} = (\mathrm{A}'_0\tan^2\beta'_2+\mathrm{E}'_0\tan^4\beta'_2)/(1+\mathrm{B}'_0\tan^2\beta'_2),$$

whence, developing $\cot^2\beta'_2$ in descending powers of M and stopping at the second term

$$\cot^2\beta'_2 = -\mathrm{A}'_0\mathrm{M}-\mathrm{B}'_0+\mathrm{E}'_0/\mathrm{A}'_0.$$

Substituting into

$$\cot^2\lambda = \cot^2\beta'_2+2\,(\mathrm{C}'_0-\mathrm{B}'_0)$$

which is valid to the same order of approximation, we obtain

$$\cot^2\lambda = -\mathrm{A}'_0\mathrm{M}+2\,(\mathrm{C}'_0-\mathrm{B}'_0)-\mathrm{B}'_0+\mathrm{E}'_0/\mathrm{A}'_0 \quad . \; . \; . \; . \; . \quad (67)$$

as the second approximation for the singular inclination when M is large, the first approximation being $\cot^2\lambda = -\mathrm{A}'_0\mathrm{M}$.

To the same order the convergency factor is

$$1+\tan^2\alpha_0\,(\mathrm{A}'_0\mathrm{M}+\{2\,(\mathrm{B}'_0-\mathrm{C}'_0)+\mathrm{B}'_0-(\mathrm{E}'_0/\mathrm{A}'_0)\}),$$

i.e.,

$$1+(n_2^2\tan^2\beta_2/n_0^2)\,\{\mathrm{A}'_0\mathrm{M}^3+(3\mathrm{B}'_0-2\mathrm{C}'_0-\mathrm{E}'_0/\mathrm{A}'_0)\,\mathrm{M}^2\}, \quad . \; . \; . \; . \quad (68)$$

β_2 referring to a ray passing through the system in the standard sense.

Now, using the equations (III), (64), (IV) and (V) of § 10 and equating suitable coefficients, we find that

$$\mathrm{A}'_0 = (n_0^2/n_2^2)\,\mathrm{A}_4, \qquad \mathrm{E}'_0 = (n_0^4/n_2^4)\,\mathrm{E}_6,$$

$$\mathrm{C}'_0 = (n_0^2/n_2^2)\,\mathrm{C}_2-4\sigma, \qquad \mathrm{B}'_0 = (n_0^2/n_2^2)\,(\mathrm{A}_3+2\mathrm{C}_2-\mathrm{B}_2)-4\sigma$$

$$3\mathrm{A}_3 = 4\mathrm{B}_2-4\mathrm{C}_2+4\,(n_2^2/n_0^2)\,\sigma, \qquad \mathrm{A}_4 = \mathrm{B}_3,$$

whence, after substitution, (68) becomes

$$1+\tan^2\beta_2\,(\mathrm{B}_3\mathrm{M}^3+\{\mathrm{B}_2-\mathrm{E}_6/\mathrm{A}_4\}\,\mathrm{M}^2). \; . \; . \; . \; . \; . \; . \; . \quad (69)$$

Now, if our B leads to a sound approximation to the convergency factor for M large, this should be

$$1+\mathrm{B}\tan^2\beta_2,$$

or, to the same approximation which we have been using

$$1+\tan^2\beta_2\,(\mathrm{B}_3\mathrm{M}^3+\mathrm{B}_2\mathrm{M}^2)\,.\quad\ldots\ldots\ldots\quad(70)$$

We see, therefore, that the development of the correct convergency factor in descending powers of M will give a result which always agrees with our B, so far as the highest term in M is concerned, but makes the term in M^2 in general different.

In the case of a lens $\mathrm{E}_6/\mathrm{A}_4$ is in general small, compared with B_2, so that this discrepancy makes little difference, but it may well be that, when we come to deal with more complicated systems, this will not be the case.

A little consideration, however, shows that, when this is so, our formula is very readily corrected so as to take this difficulty into account, without involving any lengthy numerical computation.

If we consider the formula

$$\Delta x_2 = n_2 f\,\{\mathrm{A}t_2^{\,2}+(\mathrm{E}-\mathrm{AE}_6\mathrm{M}^2/\mathrm{A}_4)\,t_2^{\,4}\}/\{1+(\mathrm{B}-\mathrm{E}_6\mathrm{M}^2/\mathrm{A}_4)\,t_2^{\,2}\}$$

it is clear that it leaves the development of Δx_2 in powers of t_2 unaltered as far as the second order inclusive. It alters the coefficient B_2 of B so as to make the two leading terms agree with (69). It also alters E in such a way as to remove the term in M^6 and reduce E to a quintic. In fact it gives for the new E the remainder obtained after the first step in the division of E by A, according to the usual process.

In practice, the terms in $(\mathrm{E}_6/\mathrm{A}_4)\,\mathrm{M}^2$ are very readily added as follows :—

$$\Delta x_2 = n_2 f\,\frac{\mathrm{A}t_2^{\,2}+\mathrm{E}t_2^{\,4}/(1-\mathrm{E}_6\mathrm{M}^2t_2^{\,2}/\mathrm{A}_4)}{1+\mathrm{B}t_2^{\,2}/(1-\mathrm{E}_6\mathrm{M}^2t_2^{\,2}/\mathrm{A}_4)}$$

and this amounts to applying the same corrective factor $1/(1-\mathrm{E}_6\mathrm{M}^2t_2^{\,2}/\mathrm{A}_4)$ or $1/(1-\mathrm{E}n_0^{\,2}t_0^{\,2}/n_2^{\,2}\mathrm{A}_4)$ to the second terms in both numerator and denominator. This factor, expressed in terms of the inclination of the incident ray, is independent of the magnification, and a short table will enable it to be found in any given case without difficulty.

A similar correction has then to be made in C; in order to keep the development of $\tan\alpha_2$ the same we must have

$$\tan\alpha_2 = t_2\,\{1+(\mathrm{B}-\mathrm{E}_6\mathrm{M}^2/\mathrm{A}_4)\,t_2^{\,2}\}\,/\,\{1+(\mathrm{C}-\mathrm{E}_6\mathrm{M}^2/\mathrm{A}_4)\,t_2^{\,3}\},$$

and writing this as

$$\frac{t_2\,\{1+\mathrm{B}t_2^{\,2}/(1-\mathrm{E}_6\mathrm{M}^2t_2^{\,2}/\mathrm{A}_4)\}}{1+\mathrm{C}t_2^{\,2}/(1-\mathrm{E}_6\mathrm{M}^2t_2^{\,2}/\mathrm{A}_4)}$$

we see that the same corrective factor has to be applied to all the second terms in the formulæ.

In the above we have used $\tan\beta_2$ as our argument, but the formulæ and the correction take precisely the same form if $\sin\gamma_2$ is the argument.

If we apply this corrective factor to the first two entries of Table III., which give a large percentage error—these correspond to cases approaching the failure of convergency and are therefore critical, we find, for lens (1)

M.	True aberration.	Formula uncorrected.	Percentage error.	Formula corrected.	Percentage error.
3	−7·87422	−7·98683	1·41	−7·90811	0·43
2	−1·52334	−1·55287	1·94	−1·54089	1·16

which shows a very sensible improvement.

The significance of this alteration is brought out more clearly when we consider the limiting case M = ∞ that is, rays actually issuing from the front focus (the case of an eye-piece). In this case, the geometrical image being at infinity, it is inconvenient to define the emergent ray by means of either longitudinal or transverse aberrations.

Let us consider the intercept of the ray on the back focal plane.

$$\text{This} = (-n_2 f\mathrm{M}+\Delta x_2)\tan\alpha_2$$

$$= n_2 f[(\mathrm{A}\tan^2\beta_2+e\tan^4\beta_2)/(1+b\tan^2\beta_2)-\mathrm{M}]\tan\alpha_2.$$

Also

$$\tan\alpha_2 = (\tan\alpha_0/\mathbf{M})(1+b\mathbf{M}^{-2}\tan^2\alpha_0)/(1+c\mathbf{M}^{-2}\tan^2\alpha_0),$$

where

$$b = \mathrm{B}-\mathrm{E}_6\mathrm{M}^2/\mathrm{A}_4,\quad c = \mathrm{C}-\mathrm{E}_6\mathrm{M}^2/\mathrm{A}_4,\quad e = \mathrm{E}-\mathrm{E}_6\mathrm{A}\mathrm{M}^2/\mathrm{A}_4,$$

and it is clear that, in the limit, where M (and $\mathbf{M}$) = ∞, $\tan\alpha_2$ must be *finite*.

This requires that b and c shall be of order M^3 and M^2 respectively, which is right, and leads to

$$\tan\alpha_2 = nb_3\tan^3\alpha_0/n\,(1+c_2\tan^2\alpha_0),$$

b_3 and c_2 being the coefficients of M^3 and M^2 in b and c respectively.

On the other hand, it is equally obvious that the intercept on the back focal plane must also approach a definite limit. Hence the factor

$$(\mathrm{A}\tan^2\beta_2+e\tan^4\beta_2)/(1+b\tan^2\beta_2)-\mathrm{M},$$

i.e.,

$$\frac{-1+(\mathrm{A}-\mathrm{M}b)\tan^2\beta_2/\mathrm{M}+e\tan^4\beta_2/\mathrm{M}}{1/\mathrm{M}+b\tan^2\beta_2/\mathrm{M}},$$

or

$$\frac{-1+(\mathrm{A}-\mathrm{M}b)\tan^2\alpha_0/\mathrm{M}\mathbf{M}^2+e\tan^4\alpha_0/\mathrm{M}\mathbf{M}^4}{1/\mathrm{M}+b\tan^2\alpha_0/\mathrm{M}\mathbf{M}^2}$$

must tend to a finite limit as M approaches ∞.

This necessarily involves that (1) A−Mb has M^3 in its leading term—a result already established, and (2) that e involves M^5 (and not M^6) in its leading term.

Hence, whenever we deal with incident rays actually passing through the front focus—and this necessarily occurs as soon as the results of the present paper are applied to aberrations off the axis—the *modified* B, C and E have to be used.

§ 15. *Combination Formulæ for More than Two Systems.*

The combination formulæ (47), (48), (51) are capable of explicit generalisation for any number of systems.

In the case of A and C successive applications of equations (47) and (51) lead at once to the results

$$\begin{aligned} f_{135\ldots 2n+1}\mathrm{A}_{135\ldots 2n+1} = {} & (f_1\mathrm{A}_1)\,\mathrm{M}^2_{35\ldots 2n+1}\mathbf{M}^2_{35\ldots 2n+1} \\ & +(f_3\mathrm{A}_3)\,\mathrm{M}^2_{5\ldots 2n+1}\mathbf{M}^2_{5\ldots 2n+1}+\cdots \\ & +(f_{2n-1}\mathrm{A}_{2n-1})\,\mathrm{M}^2_{2n+1}\mathbf{M}^2_{2n+1}+f_{2n+1}\mathrm{A}_{2n+1}. \qquad (71) \end{aligned}$$

$$\begin{aligned} \mathrm{C}_{135\ldots 2n+1} = {} & \mathrm{C}_1\mathbf{M}^2_{35\ldots 2n+1}+\mathrm{C}_3\mathbf{M}^2_{5\ldots 2n+1}+\cdots \\ & +\mathrm{C}_{2n-1}\mathbf{M}^2_{2n+1}+\mathrm{C}_{2n+1}. \qquad (72) \end{aligned}$$

in which we have reverted to the notation of §§ 2, 7.

In the case of B, we have for three systems

$$\begin{aligned} \mathrm{B}_{135} &= \mathrm{B}_{35}+\mathrm{B}_1\mathbf{M}_{35}^2+\mathrm{A}_1\mathrm{M}_{35}\mathbf{M}_{35}^2 f_1/f_{35} \\ &= \mathrm{B}_5+\mathrm{B}_3\mathbf{M}_5^2+\mathrm{B}_1\mathbf{M}_{35}^2+(f_1\mathrm{A}_1)\,\mathrm{M}_{35}\mathbf{M}_{35}^2/f_{35}+(f_3\mathrm{A}_3)\,\mathrm{M}_5\mathbf{M}_5^2/f_5, \end{aligned}$$

and the general law of combination is at once obvious.

We have

$$\begin{aligned} \mathrm{B}_{135\ldots 2n+1} = {} & \mathrm{B}_1\mathbf{M}^2_{35\ldots 2n+1}+\mathrm{B}_3\mathbf{M}^2_{5\ldots 2n+1}+\ \ldots\ +\mathrm{B}_{2n-1}\mathbf{M}^2_{2n+1}+\mathrm{B}_{2n+1} \\ & +(f_1\mathrm{A}_1)\,(\mathrm{M}\mathbf{M}^2/f)_{35\ldots 2n+1}+(f_3\mathrm{A}_3)\,(\mathrm{M}\mathbf{M}^2/f)_{5\ldots 2n+1}+\ldots+(f_{2n-1}\mathrm{A}_{2n-1})\,(\mathrm{M}\mathbf{M}^2/f)_{2n+1}. \qquad (73) \end{aligned}$$

The case of E is more difficult and we have not been able to obtain a form in which it can be written down for the combination of n systems.

But we can deal with it as follows, by determining the contribution of any one component to the whole :—

Consider three systems 1, 3, 5. It will be convenient to look upon 3 as a single lens, forming part of a larger system. 1 will then be the system of lenses preceding 3, and 5 the system following.

Applying equation (49), simplified by the use of the second invariant relation, we find

$$\begin{aligned} f_{135}\mathrm{E}_{135} = {} & f_5\mathrm{E}_5+f_{13}\mathrm{E}_{13}\mathrm{M}_5^2\mathbf{M}_5^4 \\ & +f_{13}\mathrm{A}_{13}\mathbf{M}_5^2\mathrm{M}_5^2\,(3\mathrm{A}_5/\mathrm{M}_5-3\mathrm{B}_5+4\mathrm{C}_5+4\sigma\{1-\mathbf{M}_5^2\}) \\ & +3\mathrm{B}_{13}f_5\mathrm{A}_5\mathbf{M}_5^2-2\mathrm{C}_{13}f_5\mathrm{A}_5\mathbf{M}_5^2. \end{aligned}$$

Applying the equations of combination a second time, and picking out the terms involving A_3, B_3, C_3, E_3, we find these to be

$$\begin{aligned}&f_3E_3M_5^2\mathbf{M}_5^4+M_5^2\mathbf{M}_5^4\,[f_1A_1\mathbf{M}_3^2M_3^2\,(3A_3/M_3-3B_3+4C_3)+f_3A_3\mathbf{M}_3^2\,(3B_1-2C_1)]\\&\quad+f_3A_3\mathbf{M}_5^2M_5^2\,[3A_5/M_5-3B_5+4C_5+4\sigma\,(1-\mathbf{M}_5^2)]\\&\quad+3B_3f_5A_5\mathbf{M}_5^2-2C_3f_5A_5\mathbf{M}_5^2.\end{aligned}$$

Hence the lens 3 contributes to the final E

$$\begin{aligned}&f_3A_3\{3M_5\mathbf{M}_5^2\,[A_5+(f_1/f_3)\,A_1M_{35}\mathbf{M}_{35}^2]+4\sigma\,(1-\mathbf{M}_5^2)\,\mathbf{M}_5^2M_5^2\\&\qquad\qquad\qquad+M_5^2\mathbf{M}_5^2\,[-3B_5+4C_5+(3B_1-2C_1)\,\mathbf{M}^2_{35}]\}\\&\quad+B_3\mathbf{M}_5^2\,(3f_5A_5-3f_1A_1\mathbf{M}_{35}^2M_{35}^2)\\&\quad+C_3\mathbf{M}_5^2\,(-2f_5A_5+4f_1A_1\mathbf{M}_{35}^2M_{35}^2)\\&\quad+f_3E_3M_5^2\mathbf{M}_5^4,\end{aligned}$$

and the A's, B's and C's in the curled brackets can be expressed in terms of the individual lenses of the system by means of equations (71), (72), (73).

If we denote the coefficients of A_3, B_3, C_3, E_3, in the above by l_3, m_3, p_3, q_3, then the contribution of the individual lens to E

$$= l_3A_3+m_3B_3+p_3C_3+q_3E_3.$$

Hence, if we vary K_3 for this lens, keeping focal length and magnifications unaltered

$$\Delta E = \Delta K_3\,(l_3\,\partial A_3/\partial K_3+m_3\,\partial B_3/\partial K_3+p_3\,\partial C_3/\partial K_3+q_3\,\partial E_3/\partial K_3).$$

If all the lenses are simultaneously varied, then we have

$$\Delta E = \Sigma\,\Delta K\,(l\,\partial A/\partial K+m\,\partial B/\partial K+p\,\partial C/\partial K+q\,\partial E/\partial K).$$

We have similar equations for ΔA, ΔB, ΔC, but they take a much simpler form.

Using these, we can, if we have enough lenses, vary the K's so that, *between limits*, we can make our four constants A, B, C, E take up any assigned values, or, if we wish to keep any one constant whilst slightly varying the others, we have a linear relation between the ΔK's.

§ 16. *Conclusion.*

We have now established a formula of fractional type for the longitudinal aberration of a symmetrical system which, while algebraically correct as far as the second order, does in fact, give results beyond this order in those numerical cases which have been tried, and largely overcomes the difficulties of slow convergency in critical regions.

We have further obtained a method for calculating the coefficients of this formula for any symmetrical optical system in terms of the coefficients for the components, in such a way that the effect of any single component upon the whole combination is immediately obtained.

In considering the convergency of the series usually employed, we have found that the value of the approximation depends upon the particular variable employed, and that if we wish to avoid trouble owing to lack of convergency we must use sin α_0 or tan α_0 (or a suitable multiple of these) as argument, where α_0 is the inclination of the original incident ray.

The numerical success of the new formula appears to suggest that progress in the algebraic treatment of symmetrical instruments is to be sought, not so much along the lines of developments in series, but in other mathematical directions such as continued products, or possibly continued fractions.

The next step would be to develop the method so as to cover the second order approximations to the emergent inclination. This will enable us to deal with aberrations off the axis of the system.

Some progress has already been made by the authors in this direction and the results will, it is hoped, form the subject of a later communication.

III. *Investigations on Lightning Discharges and on the Electric Field of Thunderstorms.*

By C. T. R. Wilson, *F.R.S., Reader in Electrical Meteorology in the University of Cambridge and Observer in Meteorological Physics at the Solar Physics Observatory, Cambridge.*

(*Communicated by the Director of the Solar Physics Observatory.*)

Received March 15,—Read May 6, 1920.

[Plates 2–5.]

Contents.

Section		Page
I.	Methods of measurement	74
II.	Some typical records	77
III.	On the prevailing sign of the sudden changes produced in the potential gradient by lightning flashes	85
IV.	Magnitude of the changes produced in the electric field by lightning discharges at different distances	86
V.	Effects to be expected from different kinds of discharges at different distances	87
VI.	Electric moments of the discharges	90
VII.	Quantity of electricity discharged in an average lightning discharge	91
VIII.	Electric field of a thunder-cloud	94
IX.	Conditions determining discharge	98
X.	Dimensions of the regions discharged by lightning flashes	101
XI.	Maximum potential attained before the passage of a lightning discharge	102
XII.	Mean density of the charge in a thunder-cloud immediately before a discharge	103
XIII.	Charge associated with 1 c.c. of water in a cloud	103
XIV.	Disruption of drops by the electric field	104
XV.	Pressure within a charged portion of a cloud	106
XVI.	Thunder resulting from sudden contraction due to loss of charge	106
XVII.	Energy dissipated in lightning flashes	106
XVIII.	Interpretation of "recovery" curves	107
XIX.	Electrical currents maintained in the atmosphere by thunder-clouds and shower-clouds	109
XX.	Differences between the electrical effects of shower-clouds of positive and negative polarity	110
XXI.	The normal potential gradient and air-earth current of fine weather	112
XXII.	Influence of the nature of the earth's surface below a thunder-cloud or shower-cloud	113
XXIII.	Secondary thunder-clouds	114

[Published August 27, 1920.

THE observations discussed in this paper were made at the Solar Physics Observatory, Cambridge, mainly during the summer months of 1917.

I. *Methods of Measurement.*

The method and apparatus used in the measurements are substantially those described in a paper "On Some Determinations of the Sign and Magnitude of Electric Discharges in Lightning Flashes."* The induced charge on an exposed earthed conductor (test-plate or sphere) is used as a measure of the electric field. The test-plate virtually forms part of a flat portion of the earth's surface, and the vertical electric force or potential gradient at ground level is equal (in electrostatic measure) to $4\pi Q/A$, where Q is the charge on its exposed surface and A is its area. The charge Q on the earth-connected sphere of radius R, when exposed at a height h, great compared with R, is a measure of the potential at that height; the zero potential of the sphere being the resultant of the undisturbed atmospheric potential V at the height h and of the potential Q/R due to the charge on the sphere, so that $Q/R = -V$. The earthed conductors can be shielded from the earth's field: the test-plate by means of an earth-connected cover, the sphere by lowering it into a conducting case resting on the ground. The quantity of electricity which flows to earth through the connecting wire on exposing or shielding the test-plate or sphere, is measured by a special type of capillary electrometer in which the readings indicate the total quantity of electricity which has traversed the instrument; the sign and magnitude of the charge on the exposed conductor, and thus of the potential gradient, at the beginning and end of an exposure are thus determined. The sign and magnitude of sudden changes of potential gradient which occur while the conductor is exposed are indicated by the direction and magnitude of the resulting displacements of the electrometer meniscus. The total flow of electricity between the atmosphere and the test-plate or sphere during an exposure is also measured —being given by the difference between the electrometer readings before and after the exposure. The principal improvement introduced has been the provision of apparatus for giving a photographic trace of the electrometer readings; rapid changes in the field occupying less than one-tenth of a second are in this way recorded.

In the observations described in the previous paper the sphere was supported in a manner which did not admit of absolute measurements being made, as the charge measured included that on the upper part of the support as well as that on the sphere itself; in these earlier measurements therefore the sphere was standardised by comparison with the test-plate. The method of supporting the sphere is now such that the charge on the sphere alone is measured, while the disturbing effect of the earthed supporting rod is small, and thus the potential at the level of the earth-connected sphere can be calculated from the charge upon it. The new method of mounting the sphere is shown in fig. 1.

---

* 'Roy. Soc. Proc.,' A, vol. 92, p. 555, 1916.

The sphere, 30 cm. in diameter, is supported on an earth-connected brass tube B, 2 cm. in diameter, from which it is insulated by sulphur-coated ebonite E; insulators are indicated in the figure by the dotted areas. The tube is inserted within a wider one C which extends from the top to the bottom of the sphere and which is open below. The supporting tube B is rigidly fixed in a hole bored through the screw cap which closes the upper end of an iron pipe P, 5 cm. in external diameter and 427 cm. long, which can be turned about its lower end from the vertical to a nearly horizontal position as described in the former paper. The length of the brass tube from the top of the iron pipe to the bottom of the sphere is 38 cm. The connection between the sphere and the electrometer is made by means of a tightly stretched wire W supported by quartz insulators. The wire is not attached directly to the sphere but to a brass disc D insulated from the supporting tube and fitting loosely within the wider tube C inside the sphere. The sphere is fixed to the disc by means of a screw which projects from its inner surface and can thus readily be removed to give access to the insulation.

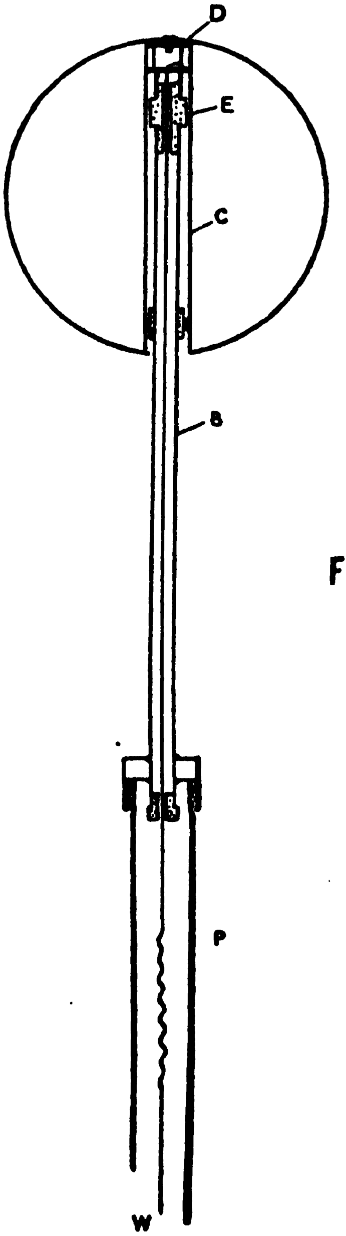


FIG 1

When the sphere is exposed by raising the iron pipe to its vertical position the height of its centre above the ground is 480 cm.

The sphere when lowered is received in a metal-lined earthed box resting on the ground; a tightly fitting cover, also metal lined and earthed, protects the sphere from the atmospheric electrical field and from the weather. The charge on the earthed sphere in this position is taken as zero.

The charge Q on an earthed sphere of radius R at a height h above level ground is assumed to be such that $Q/R-Q/2h+V=0$, where V is the undisturbed air potential at the height h. The presence of the neighbouring hut exerts a disturbing influence which however is not large: the correction to be applied has been estimated by imagining the hut to be replaced by a conducting hemisphere large enough to enclose it. The vertical potential gradient over level ground being assumed uniform throughout a height exceeding that of the hemisphere, the lowering of potential at a given point by the induced charge on the hemisphere is readily seen to be equal to Va^3/r^3 where a is the radius of the hemisphere and r is the distance of its centre from the given point. In the actual case the correction amounts to 6 per cent.

The charge on the exposed earthed test-plate (the surface of which is at ground level) is similarly diminished by the presence of the hut; the correction to be applied amounts in this case to about 1 per cent. A somewhat larger correction—estimated at 1·5 per cent.—has to be made for the effect of the induced charge of the earth-connected cover and its supporting arm. Apart from these small corrections the relation between the potential gradient F at ground level and the charge Q on the exposed earthed test-plate, of area A, is given by $4\pi Q/A = F$, when the quantities are expressed in C.G.S. electrostatic measure. The effective area of the plate is 2220 sq. cm.

For the measurement of the quantities of electricity which passed between the exposed conductor and earth through the connecting wire, the capillary electrometer described in the previous paper was used. By means of a $\frac{1}{4}$-inch microscope objective, placed with its axis vertical above the electrometer, an image of the meniscus was formed on a horizontal slit. The slit coincided in position with the image of the axis of the capillary tube and was almost in contact with the sensitive surface of a photographic plate kept in uniform motion at right angles to the slit. It was made by ruling a line with a razor blade on an exposed and developed "process" plate; it was protected by a strip of microscope cover-glass cemented with Canada balsam to the gelatine surface—the thin cover glass was next the moving photographic plate and was only a small fraction of a millimetre distant from it. The breadth of the slit was about $\frac{1}{50}$ mm.

The carrier of the photographic plate was clamped to the middle portion of a wire stretched horizontally over two pulleys; a weight was attached to one end of the wire, while the other was attached to a piston, the motion of which in its cylinder caused oil to be driven through a fine hole in a brass disc. By turning the disc any one of a graduated series of holes could be brought into action according to the speed of travel desired.

The light from the source of illumination—a paraffin lamp—could be cut off momentarily by means of a shutter which was worked by a cord from outside the hut. In this way it was possible to record on the photographic plate the times of the beginning and ending of thunder. In the records reproduced (Plates 2 to 5) these momentary interruptions of the illumination are represented by vertical black lines; a single line indicates the beginning, a double line the end of a peal of thunder.

The interval on the photographic record between the vertical portion of the trace, which represents the sudden change of field due to the passage of a lightning flash, and the dark line which marks the moment when the thunder resulting from the flash began to be heard, afford data for obtaining an estimate of the approximate distance of the discharge.

The varying position upon the slit of the image of the meniscus on which the microscope is focussed is represented by the curve separating the dark and light regions of the record. The fine horizontal lines are due to dust particles or to irregularities of the slit; they are useful as reference lines from which the displacement of the meniscus may be measured. The vertical flutings which appear in some of the records are probably due to flickering of the lamp.

Records of the electrical effects of thunderstorms at various distances from the place of observation were obtained on ten different days in 1917. The records were not by any means continuous throughout the whole duration of a storm: comparatively quick runs of the recording apparatus were generally made—varying from 3 to 50 minutes in duration—and some time had to be spent in changing the photographic plates and readjusting the apparatus between the successive runs. Again, the difficulty of estimating the order of magnitude of the electric effects to be expected frequently led to the sphere being exposed when the test-plate would have been more suitable, or *vice versa*; the readings—which are 40 times larger with the sphere than with the test-plate—being in consequence too large or too small to be recorded. Thus the records obtained served rather to sample a thunderstorm at different stages of its history than to give a complete account of the changes in its electric field.

II. *Some Typical Records.*

Enlargements of some of the records are reproduced in Plates 2 to 5. In the original negatives a change of one mm. in the ordinates represented a flow of 24 electrostatic units, or $8 \cdot 0 \times 10^{-9}$ coulomb through the electrometer: a change of potential gradient of 100 or 4000 volts per metre, according as the sphere or the test-plate was used, was required to cause the passage of this quantity of electricity between the exposed conductor and the earth.

A fairly typical fine weather record (May 23, 1917, 14h. 17m. to 14h. 51m. G.M.T.) is shown in fig. 1, Plate 2: the sphere was used as the exposed conductor. The record begins with a horizontal portion traced before the conductor was exposed to the electric field. The small peak near the beginning of the record shows the effect of raising the sphere to its maximum height (480 cm.), and immediately lowering it again into its protecting case; it indicates the existence of a positive potential gradient of 100 volts per metre. The sphere was raised at 14h. 20m., the exposure to the electric field being continued till 14h. 50m. except for regular interruptions at 5-minute intervals when the sphere was momentarily lowered into its case. The depths of the notches in the curve are measures of the potential gradient at the times of lowering the sphere: the potential gradients recorded at intervals of 5 minutes are, in volts per metre, 120 (at 14h. 30m. when the sphere was raised), 110, 120, 90, 75, 80, and finally 90 at 17h. 50m. (when the sphere was lowered).

The difference in the ordinates of the final and initial horizontal portions of the trace (both recorded while the sphere was in its case) is a measure of the integrated ionization current which has entered the sphere during the 30 minutes' exposure to the atmospheric electrical field. The record shows that this amounted to 21·8 E.S.U., while the mean charge induced on the earth-connected sphere during the exposure amounted to 23 E.S.U.—the equivalent of 96 volts per metre. The mean "dissipation factor" for the period of exposure was thus $\frac{21 \cdot 8}{30} \times \frac{100}{23}$, *i.e.*, about 3 per cent. per minute.

The readings obtained when the sphere is down form a series of points on a curve of which the vertical height above the initial horizontal part of the trace is a measure of the integrated ionization current which has entered the sphere from the atmosphere. This curve forms the zero line for potential gradient, *i.e.*, the differences of the ordinates of this curve and of the actual trace obtained when the sphere is exposed give a measure of the potential gradient at any moment.

In fig. 2 is reproduced the record of May 12, 1917, from 16h. 50m. to 17h. 35m. The sky was overcast and the weather conditions suggested thunder—a storm did in fact occur some hours later. The sphere was momentarily raised at 16h. 51m.; raised again at 16h. 55m., and kept up till 17h. 23m., being however momentarily lowered into its case at 5-minute intervals during this time; it was kept in its case after 17h. 23m. The potential gradient was 150 volts per metre at 16h. 51m.; it gradually diminished till it reached negative values, and continued to be negative from 17h. 12m. 50s. till 17h. 18m. 10s., reaching a minimum of −80 volts per metre at 17h. 16m., becoming positive again and being equal to 260 volts per metre when the sphere was lowered at 17h. 23m. The negative potential gradient coincided in time with the passage overhead of a cloud discharging rain which did not reach the ground. The test-plate was uncovered from 17h. 25m. to 17h. 30m.: the displacement of the meniscus on uncovering and covering the plate is almost too small to be seen in the reproduction of the record but indicates the continuance of a positive potential gradient of about 300 volts per metre. The ionization current from the earth-connected sphere to the atmosphere during the period of negative potential gradient has been sufficient to neutralise approximately the flow from the atmosphere to the negatively charged earth-connected sphere during its exposure to the positive potential gradient.

All the remaining records reproduced in the plates show the effects of lightning discharges (generally at a considerable distance) on the potential gradient.

Fig. 3 (June 13, 1917, 14h. 11m. to 14h. 16m. 30s.).

The sphere was exposed during the whole time represented by the record except at about 14h. 12m. 30s., when it was momentarily lowered; the effect of lowering and raising the sphere is indicated by the prominence midway between 14h. 12m. and 14h. 13m. The potential gradient at that moment was negative and equal to −420 volts per metre. The summit of the prominence gives the zero line of potential gradient. The record begins with a negative potential gradient of about −430 volts er metre. At 14h. 11m. 10s. distant electrical charges which were responsible for a ortion (amounting to 150 volts per metre) of the negative potential gradient at the lace of observation were neutralised by the passage of a lightning flash. The negative otential gradient at once began to be regenerated but was again suddenly diminished bout 3 seconds later, losing 25 volts per metre by the passage of a lightning flash, robably at a still greater distance. This continuous production of a negative potential radient and its sudden diminution at intervals by lightning discharges continues hroughout the record. At about 14h. 13m. 40s. a sudden change of potential gradient f positive sign occurred, but was followed by one of negative sign and of nearly qual magnitude about 0·4 second later, a small positive change again occurring after nother almost equal interval; these changes of potential gradient amount to +240, −220 and +25 volts per metre respectively. Another negative change of potential radient (about 60 volts per metre) is indicated 10 seconds later. A few seconds fter 14h. 16m. the record shows two discharges to have occurred with an interval of 2·4 seconds between them; each produced a change of potential gradient of positive ign, the first amounting to 840, the second to 870 volts per metre.

The potential gradient at any moment may be regarded as being the resultant of everal electric fields, including those due to charges concentrated in different thunder-louds or different centres of activity in the same cloud. The passage of a lightning lash results in the sudden destruction of one of these constituent fields. This at nce begins to be regenerated by processes going on in the thunder-cloud at a rate which is indicated by the slope of the curve. The curve of recovery of the electric ield (approximately logarithmic) shown after the discharges of 14h. 14m. is quite typical; similar curves appear in most of the records, a specially striking example being that of fig. 11 (Plate 4).

On account of the very short intervals between the successive peals of thunder, the times at which they began and ceased to be heard were not systematically recorded during the record reproduced in fig. 3. The first peal of thunder recorded is marked by the single and double dark lines as beginning at 14h. 13m. 8·9s. and ending at 14h. 13m. 15s., and a second one as beginning at 14h. 13m. 18·4s. and ending at 14h. 13m. 30s. The two peals are taken as being due to the discharges at 14h. 12m. 47·7s. and 14h. 12m. 53·6s. respectively; this gives a distance of 7·1 km. for the first and of 8·3 km. for the second. The first of these discharges produced a total change of potential gradient of 350 volts per metre, but this took place in two stages of 220 and 130 volts per metre which were separated by an interval of about 0·2 seconds; this interval is barely distinguishable in the reproduction. The discharge at 8·3 km. produced a change of about 95 volts per metre. The peal of thunder marked as beginning at 14h. 14m. 9s. probably belongs not to the flash at 14h. 14m., but to an earlier one, possibly the double one at 14h. 13m. 40s.

Fig. 4 (August 9, 1917, 14h. 45m. to 15h. 2m. 30s.).

Here the test-plate was exposed in place of the sphere. The potential gradient indicated at 14h. 45m. 15s. when the plate was first uncovered is negative (−4570 volts per metre). The principal sudden changes of potential gradient (all in the neighbourhood of 3000 volts per metre) are negative, indicating the destruction of positive fields by the passage of lightning discharges. The times of the beginning and ending of the peals of thunder were in most cases marked as shown by the single and double black lines. The distances indicated by the intervals between the principal discharges and the beginning of thunder are all about 5 km. The characteristic curve of recovery after the passage of each discharge is well shown.

Heavy black clouds were overhead at the beginning of the record and slight rain began about 14h. 48m. 30s. and became heavy at about the time when the record ceased. The effect of the rain is shown by the downward slope of the latter portion of the trace, which indicates a flow of positive electricity from the earth through the capillary electrometer to the test-plate. How much of this positive charge went to increase the induced positive charge on the test-plate (on account of increasing negative potential gradient) and how much to neutralise a negative charge carried down to the test-plate by rain drops, or by ions travelling under the influence of the negative potential gradient, remains undetermined owing to the fact that the cover was not replaced until after the record was completed.

Fig. 5 (June 17, 1917, 20h. 23m. 23s. to 20h. 27m. 29s.).

This is an enlargement of a portion of a record obtained while a severe storm was passing at a distance of 15 to 20 km. Between 20h. 20m. and 20h. 29m. the photographic trace recorded 95 positive discharges (*i.e.*, discharges causing a sudden positive change of potential gradient) and 40 negative discharges. The discharges were visible as vertical flashes passing between a cloud near the N.W. horizon and the earth, many of the flashes being multiple. The storm was seen to travel from W. to N.; newspaper reports show that it passed over St. Ives, which lies from 10 to 11 miles (about 17 km.) to the N.W., damage by lightning occurring there. The mean of the 95 sudden changes of potential gradient of positive sign amounted to 119 volts per metre, that of the 40 of negative sign to −80 volts per metre.

The sphere was used in obtaining this record.

Fig. 6 (June 16, 1917, 19h. 12m. to 19h. 23m.).

The test-plate was used as the exposed conductor.

The potential gradient was negative (−5400 volts per metre) at 19h. 12m. 45s. when the cover was removed, positive (=1000 volts per metre) at 19h. 22m. 15s. when the cover was replaced. Rain was falling throughout the duration of the record, and the charge carried down (by rain and ionization current) during the 9½ minutes' exposure was negative and amounted to 16×10^{-12} coulombs per sq. cm., the mean current being thus about 27×10^{-15} ampere per sq. cm. Two of the discharges recorded—at 19h. 17m. 4s. and at 19h. 20m. 55s.—were multiple, as is shown in the enlargements, figs. 18 and 19 of Plate 5. All the sudden changes of potential were negative—excepting the positive components of the multiple flashes—the largest amounting to −9600 volts per metre. The distance of this discharge, as is shown by the interval elapsing before the thunder began to be heard, was about 4·3 km. The distances of the others ranged between 4·3 and 5·7 km. The peals of thunder, as the intervals between the single and the double black lines show, were very long, some lasting for as many as 40 seconds.

Fig. 7 (June 16, 1917, 19h. 31m. 10s. to 19h. 36m. 45s.).

This is a portion of the record taken with the test-plate next after that shown in fig. 6. Rain continued to fall throughout the time of exposure. The cover was removed from the test-plate at 19h. 31m. 30s.; the potential gradient at that moment was negative, −6500 volts per metre; a lightning discharge had probably occurred immediately before the exposure of the test-plate. The discharge at 9h. 33m. 20s. was really multiple, the sudden changes of potential gradient being −2900, −5100 and +1300 volts per metre. The discharge at 19h. 35m. 55s. was negative (change of potential gradient = −4100 volts per metre) and was at a distance of about 5·5 km. The characteristic form of the recovery curve following the discharges is modified by the superposition of a general downward slope which represents an electric current from the ground to the atmosphere; this current was probably mainly carried by falling negatively charged raindrops. The potential gradient when the cover was replaced at 19h. 39m. (beyond the limits of the portion of the record reproduced) was positive and exceeded 2500 volts. The total quantity of electricity transferred per sq. cm. of the test-plate to the atmosphere, during the whole 8½ minutes of exposure, but mainly after 19h. 36m. was the equivalent of 40,000 volts per metre, *i.e.*, $3{\cdot}5 \times 10^{-11}$ coulomb.

Fig. 8 (May 29, 1917, 19h. 4m. 10s. to 19h. 11m. 50s.).

This is the final portion of a record which began at 18h. 44m. The sphere was used as the exposed conductor. The potential gradient had been +1200 volts per metre at 18h. 46m. when the sphere was raised; 980 at 18h. 51m. and −1400 at 18h. 56m., at which times the sphere was momentarily lowered. The peak shown in the figure at 19h. 6m. represents the effect of again momentarily lowering the sphere and indicates that the gradient was still negative, being equal to −430 volts per metre. At 19h. 11m. when the sphere was finally lowered the potential gradient had again become positive, being now +20 volts per metre.

These comparatively gradual changes of potential gradient accompanied the passage of towering cumulus clouds at no great distance. Superimposed upon them are sudden changes (amounting at most to 150 volts per metre) produced in the field by frequent discharges of more distant thunder-clouds. Some of these are positive, some negative; the discharges of either sign are alike in being followed by the characteristic curve of recovery of the field.

Fig. 9 (May 29, 1917, 17h. 58m. 45s. to 18h. 12m. 30s.).

This is an enlargement of a portion of a record which extended from 17h. 56m. to

18h. 36m.; the sphere was exposed. During the period covered by the portion of the record reproduced the sphere was momentarily lowered at 18h. and at 18h. 5m; it was also lowered at 18h. 10m. and kept in its case till 18h. 11m. when it was again raised. The potential gradient at the times of lowering the sphere amounted to +90, +60 and +40 volts per metre. It is plain from the record that the gradient remained positive throughout: the principal sudden changes of gradient were positive, and amounted to about 150 volts per metre; two, however, at about 18h. 1m. and 18h. 9m. 30s. were negative and equal to about 60 volts per metre. Positive discharges evidently also occurred during both the short periods for which the sphere was lowered.

The characteristic recovery of the field after both positive and negative discharges is well shown. The two peals of thunder recorded probably belong not to the discharges immediately preceding them but to the previous discharges. The discharges were probably at a distance of 20 km. or more.

Fig. 10 (August 15, 14h. 18m. 20s. to 14h. 30m. 15s.).

At 14h. 18m. 30s., when the cover was removed, the potential gradient was negative (= −3600 volts per metre). This negative potential gradient had increased to about −5000 at 14h. 19m. 6s.; at this moment the negative field was nearly destroyed by the passage of a lightning flash at a distance of 4·1 km., the sudden change in potential gradient being +4800. The five subsequent flashes also produced positive changes in the potential gradient; the beginning and ending of the thunder is marked on the record in each case. The magnitudes of the sudden changes of potential gradient vary from 3900 (the third shown in the fig.) to 14,600 volts per metre (the last); the distances of these two discharges were practically the same, 3·7 and 3·8 km.

The striking feature of this record is the abnormal character of the curve of recovery of the field after the passage of every discharge except the last; instead of the rate of recovery of the field being most rapid immediately after the discharge, it is at first zero or very small and gradually increases to a maximum, falling off again with the increasing field as in the normal type. The last discharge shown in fig. 10 as well as all the subsequent discharges of the record of which this is a part were followed by a recovery of the field of the normal type. The recovery curves following the discharges of the immediately preceding record of the same storm were also normal in character.

Rain began about 14h. 20m., became heavy about 14h. 25m., and ceased about 14h. 31m. 30s.

The potential gradient was negative throughout the period covered by fig. 10 until reversed by the last discharge shown. At 14h. 28m., when the cover was momentarily replaced, the potential gradient was −4800 volts per metre. The small hump in the curve at 14h. 24m. 45s. is due to the shielding effect of a horse and cart which passed within a few yards of the test-plate.

In spite of the negative potential gradient, which would tend to produce an ionization current from the ground to the atmosphere, the total charge received by the test-plate from the atmosphere between 14h. 18m. 30s. and 14h. 28m. has been positive and equal to the charge which a potential gradient of 17,000 volts per metre would have induced on the earth-connected plate. The charge carried by the rain must thus have been positive and must have exceeded to the above extent the negative charge carried by the ionization current. The greater part of this charged rain has evidently fallen between 14h. 26m. and 14h. 28m.

Fig. 11 (June 12, 1917, 16h. 38m. 40s. to 16h. 50m.).

This is a portion of the second record taken on a sultry afternoon with towering cumulus in all directions. The first record ran from 15h. 55m. to 16h. 19m. A cap was seen to form on the summit of a large cumulus cloud in the E.N.E. at 15h. 59m., and another on one of the lower heads of the same cloud now in the N.E., about 16h. 16m. The potential gradient throughout this first record was positive and about 50 volts per metre. No thunder was heard and no sudden changes of the field are shown on the record.

The second record, of which fig. 11 is a portion enlarged, extended from 16h. 27m. to 17h. 4m. The large cumulus cloud was N. by E. with its edge at an elevation of about 60 degrees at 16h. 30m. and due N. about 16h. 50m. The potential gradient diminished from +44 volts per metre at 16h. 30m. to 15 at 16h. 35m. and then became negative, being −29 at 16h. 40m., −175 at 16h. 45m. (all the above being occasions of momentary lowering of the sphere), and −106 volts per metre at 16h. 49m. when the sphere was finally lowered. The field was zero at 17h. 14m., and had become positive at 17h. 16m., when the observations ceased.

All the sudden changes of field observed were positive; two of 18 and 14 volts per metre, both due to flashes at a distance of about 7 km., occurred before the field had become negative. The other two (equal to 120 and 320 volts per metre respectively) are shown in the figure; they both show characteristic recovery curves; in both cases the potential gradient was reversed, in the first by a discharge at a distance of 8·2, in the second by one at a distance of about 7 km.

The discharge at 16h. 45m. 50s. is an interesting one. The negative potential gradient had reached a steady value—about 170 volts per metre—before the passage of the discharge. The discharge—at a distance of about 7 km.—caused the gradient to become positive (=150 volts per metre); the negative field was again re-established, practically exponentially, a steady value being finally reached equal to about 105 volts per metre. The sphere was brought down at 16h. 49m. No thunder was heard after this time.

In Plate 5 are enlargements of small portions of some of the traces, showing details in the changes of potential gradient associated with lightning discharges. In each case the time in seconds is shown, reckoned from the moment at which the discharge, as indicated by the record, began.

The principal discharge of fig. 12 occurred on May 29, 1917, at about 15h. 23m. 10s.; the peal of thunder which followed began 21·5s. later (indicating a distance of about 7 km.) and was audible for about 20 seconds. The sudden change produced in the potential gradient was negative and exceeded 1250 volts per metre. The record shows the characteristic curve of recovery of the field, interrupted at 1m. 50s. after the discharge by the lowering of the sphere. The positive field due to the charge which the flash neutralised was nearly counterbalanced at the place of observation by a negative field, so that the resultant potential gradient before the passage of the discharge was only about +360 volts per metre.

When the sphere was first raised, at 15h. 19m. 30s., the potential gradient was positive—about 150 volts per metre—and it increased up to the moment of the principal discharge. There were, however, during this time, small sudden changes of potential, some positive, others negative, none exceeding 50 volts per metre; they were obviously due to very distant discharges; no thunder was recorded. Throughout the afternoon there were towering cumulus clouds in all directions, rain falling from some of them.

The discharge of fig. 13 occurred about 15h. 9m. 30s. on August 15, 1917. The sphere was lowered at 15h. 10m., the characteristic curve of recovery of the field being thereby interrupted. The peal of thunder began while the sphere was being lowered, *i.e.*, about 40 seconds after the discharge; the beginning is not marked on the record, but the double dark line indicates that the peal of thunder ended about 55 seconds after the discharge. The potential gradient immediately before the discharge had a negative value exceeding 1000 volts per metre; immediately after the discharge the potential was positive and equal to about 300 volts per metre.

The first discharge of fig. 14 occurred at 13h. 50m. on June 13, 1917, just at the moment when the sphere had been raised to its exposed position. The potential gradient before the discharge was negative (= − 690 volts per metre). The discharge was a double one, causing an increase in the negative potential gradient of more than 980 volts per metre, followed by a sudden change of the opposite sign, which brought the potential gradient to within 260 volts per metre of its original value, the total duration of the double discharge being about one-fifth of a second.

Two other double discharges of about the same total duration were recorded about 22 seconds and 87 seconds later, the first giving sudden potential changes of +70 and −30, the second of +100 and −115 volts per metre. The other discharges shown in the figure are noteworthy as not being followed by the usual recovery curve.

The next three figures are further examples of double discharge records of the same type—*i.e.*, of records showing the occurrence within a very short interval of time of two sudden changes of potential of opposite sign. They differ among themselves mainly in the relative magnitudes of the two sudden changes: the first change of gradient being the greater in fig. 15, the two being approximately equal in fig. 16, and the second being the greater in fig. 17; in the last case the initial change is negative, in the two others positive. The duration of the double discharge is about one-fifth of a second in fig. 15, two-fifths of a second in fig. 16, and two-fifths of a second in fig. 17.

Double discharges consisting of two successive sudden changes of potential gradient of the same sign are also not uncommon. A striking example is that of the last discharge shown in fig. 3, where a sudden positive change of potential of 840 volts per metre is followed 2·4 seconds later by a second change of the same sign amounting to 870 volts per metre. The discharge at 14h. 12m. 50s. (in the same fig. 3) was also really a double one of this type, the interval between the two components of magnitudes, 220 and 130 volts per metre, being about one-fifth of a second.

What have been called above double discharges, it should be noted, are not necessarily discharges along the same track or even from the same thunder-cloud; it may often be observed that lightning flashes from two different centres occur almost simultaneously.

In the last three figures of Plate 5 are reproduced enlargements of records of multiple discharges, *i.e.*, of records showing a rapid succession of changes of potential gradient of opposite sign. These were all obtained during the same thunderstorm, that of the afternoon of June 16, 1917. The first shows sudden changes of potential gradient of −9600, +4350 and −1500 volts per metre, the intervals between the reversals being about one-third of a second. The second shows sudden changes of potential gradiant amounting to −7100, +1700, −1700, +300, −1900, +700, −600, +1000, the total time occupied by the eight reversals being 2·1 seconds. In the third the changes of potential gradient are −1600, +900, −1600, +1200, +700, −700 volts per metre, the total duration being 1·9 seconds.

III. *On the Prevailing Sign of the Sudden Changes Produced in the Potential Gradient by Lightning Flashes.*

The sudden changes produced in the potential gradient at the place of observation by the passage of lightning discharges have been more often positive than negative, *i.e.*, the greater number have consisted in a sudden increase of a previously existing positive potential gradient or a diminution or reversal of a previously existing negative gradient; in other words they might be explained as being due to the discharge of a negatively charged cloud. Discharges producing such a positive change of potential gradient are called in what follows positive discharges.

The number of positive discharges recorded in 1917 was 432, of negative discharges 279. If the observations of 1914 and 1915 are included, the numbers are 528 and 336, the ratio being 1·56. Of the ten days of thunder on which records were obtained in 1917, there were nine on which more positive than negative discharges were recorded; on the remaining day, however (June 16) about twice as many negative as positive discharges were recorded (74 negative, 38 positive). It is perhaps natural to associate the excess of positive over negative discharges with the excess of positive electricity found by SIMPSON* and others to be carried down in rain, the greater part of the charge transferred from the atmosphere to the earth by the rain of the thunderstorm being perhaps returned in lightning discharges. (See however Sections XIX. and XX.)

IV. *Magnitude of the Changes Produced in the Electric Field by Lightning Discharges at Different Distances.*

The approximate distance of each lightning flash which caused a disturbance on the photographic trace was, when possible, determined by observing the time interval between the discharge and the thunder associated with it. The beginning of each

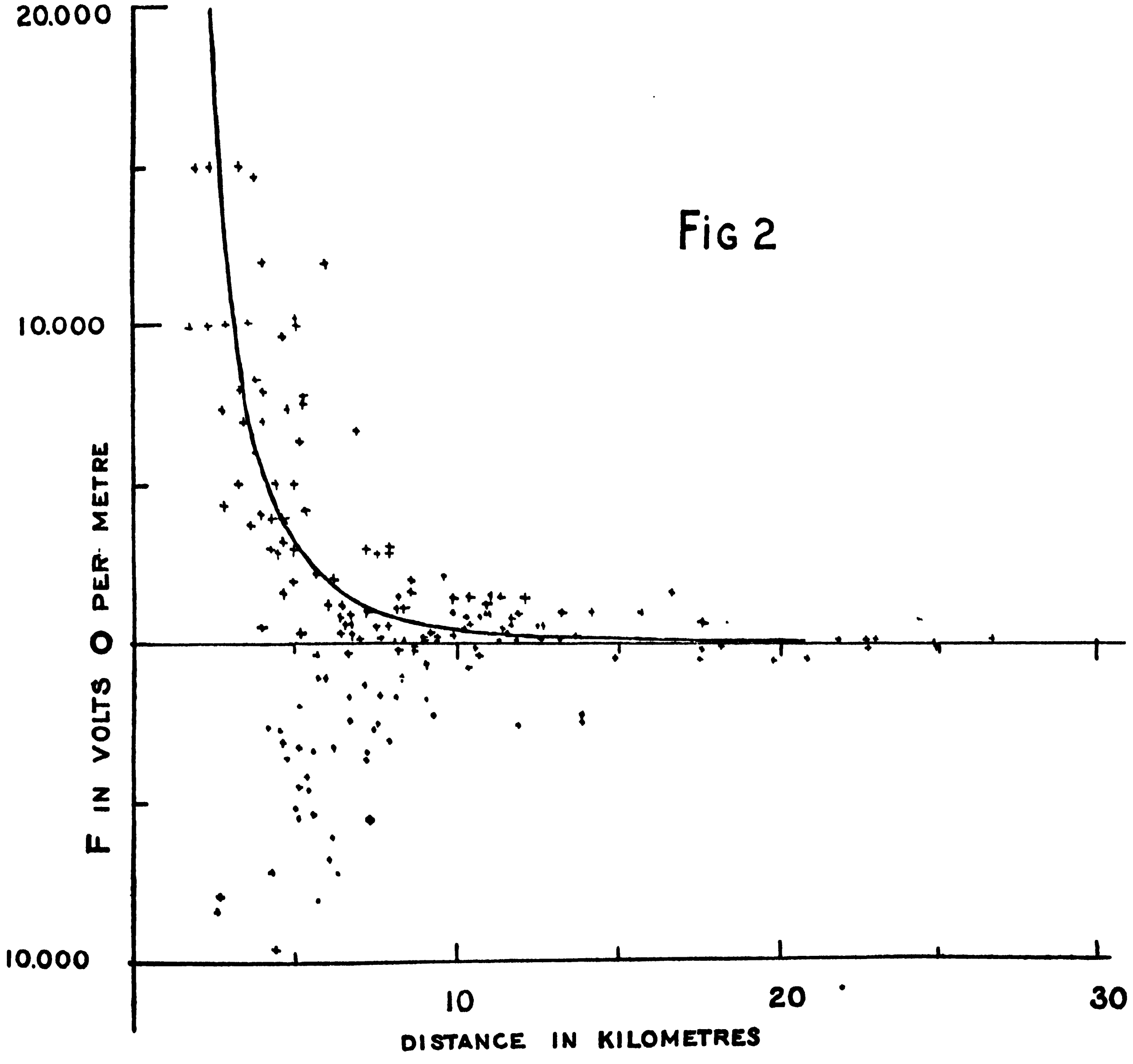


FiG 2

---

* SIMPSON, 'Phil. Trans.,' A, vol. 209, p. 379, 1909.

peal of thunder heard was marked on the trace by momentarily cutting off the light as described in Section I. It was by no means always possible to be certain which peal of thunder recorded was caused by the lightning discharge responsible for a given sudden disturbance of the field; when the storm was a distant one with very frequent lightning flashes there might be several subsequent discharges between the passage of a flash and the arrival of the sound of its thunder. There appeared to be no ambiguity in the case of about 120 discharges recorded in 1917; the approximate distance L of each of these discharges and the sign and magnitude of the resulting sudden change of field F are shown in fig. 2, which includes also the eye observations of 1914 and 1915. When the records show two or more sudden changes of field within a fraction of a second it is the largest of these which is recorded in fig. 2; it was considered that if the component discharges of the multiple flash were not all at the same distance, the one which produced the largest effect was likely to be the nearest, and therefore that of which the distance was deduced from the interval elapsing between the discharge and the beginning of the thunder.

V. *Effects to be Expected from Different Kinds of Discharges at Different Distances.*

A lightning flash may consist in the passage of a charge Q from a certain region A of the atmosphere to earth, or from a region A_1 to another A_2 both in the atmosphere; A_1 and A_2 may be in the upper and lower parts of the same thundercloud with their centres near the same vertical line, or they may be at a considerable horizontal distance apart.

Let a charge Q derived from a certain region A of the atmosphere pass to earth. The change in the electric field may be considered as due to the removal of the charge Q from A and of an equal and opposite charge $-Q$ from A' the image of A. Just as for many purposes no sensible error is made by assuming the magnetism of a bar magnet to be concentrated at two definite points, the poles of the magnet, so in the present case the charges Q and $-Q$ may be regarded as being concentrated at two points p and p'. These points are situated at a height H above and at an equal depth below the surface of the ground, such that $2QH = 2\Sigma qh = M$, the electric moment of the discharge;* q being the charge derived from a small element of volume at a height h. In calculating the change produced in the electric field at distant points by the passage of the discharge, no sensible error will be made by making this substitution, and even at points at no great distance from the axis pp', the error will be small if there has been any approximation to a symmetrical distribution of the charge in a sphere surrounding p.

---

* In the present paper $2qh$, not qh as in the previous paper, is taken as the electric moment of the discharge of a quantity q from a height h to earth, the moment with which we are concerned being that of charge q at a height h together with that of its image $-q$ at a height $-h$.

The vertical electric force at a point on the ground is given by

$$F = \frac{2QH}{L^3\left(1+\frac{H^2}{L^2}\right)^{\frac{3}{2}}} = \frac{2Q}{H^2\left(1+\frac{L^2}{H^2}\right)^{\frac{3}{2}}}$$

$\left(\text{becoming } \frac{2QH}{L^3} \text{ and } \frac{2Q}{H^2} \text{ when L is large and small respectively compared with H}\right)$ where L is the distance of the point from the axis *pp'*. The curve I in fig. 3 represents

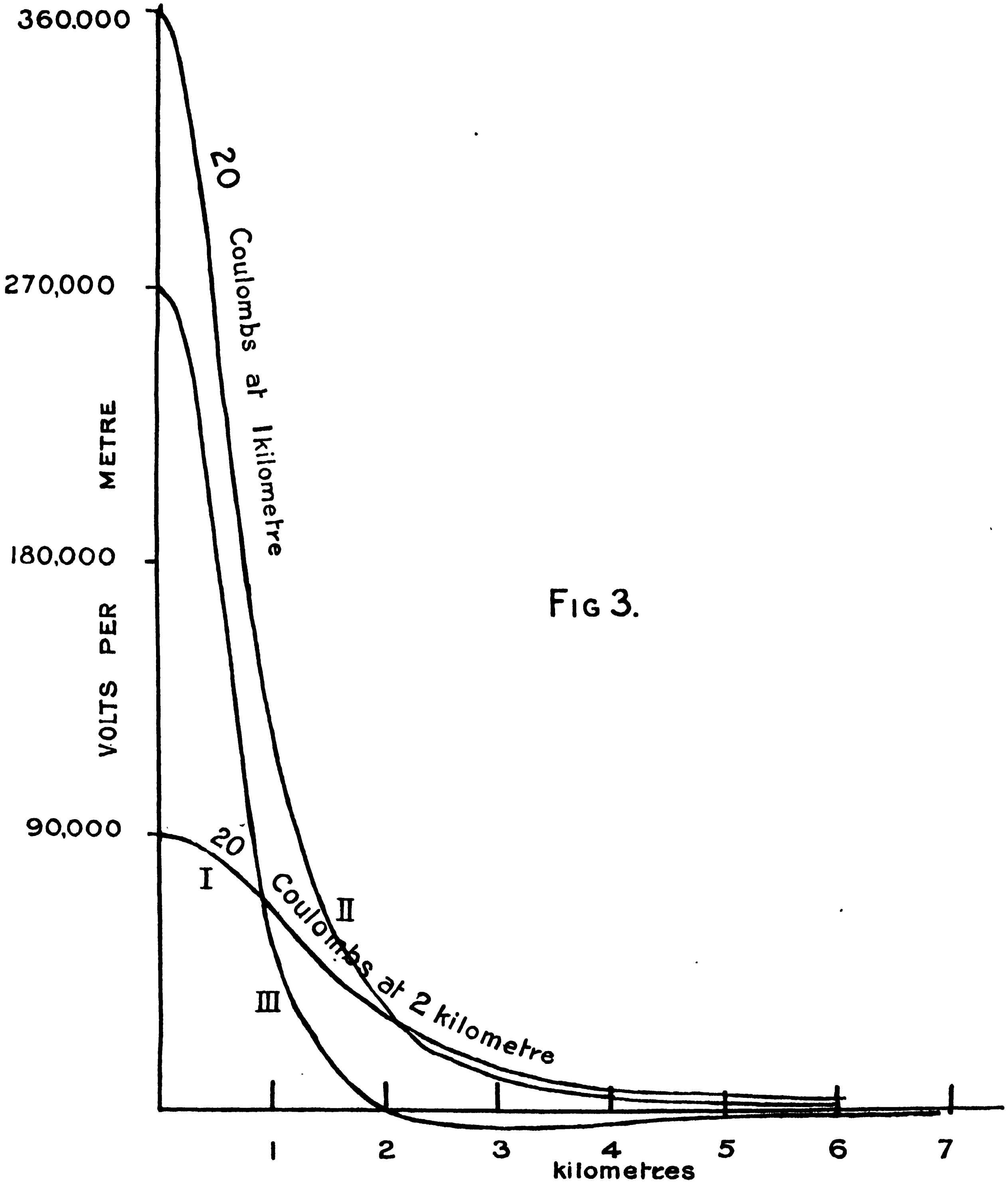


FIG 3.

the potential gradient at the earth's surface at different distances due to a charge Q supposed concentrated at a point at a height of 2 km.: it represents the change produced in the potential gradient by the discharge of 20 coulombs from a height of 2 km. to earth. The curve II. represents similarly the potential gradients due to the same charge at a height of only 1 km. The difference between the ordinates of the two curves (curve III.), represents the change of potential gradient produced by a vertical discharge of the quantity Q from a height of 2 km. to a height of 1 km. The sign of the effect is reversed at a certain distance, the vertical electric force at the surface of the ground being, for a given charge Q at the lower level, greater than for the same charge at the higher level when the distance is small, but less when the distance is great.

For a vertical discharge from a height H_2 to a height H_1, F, when L is large, becomes equal to $2Q(H_2-H_1)/L^3$, so that $FL^3 = 2Q(H_2-H_1)$ which may be defined as the electric moment of such a discharge. Thus when L is large FL^3 is equal to the electric moment of the discharge whether this reaches the earth or not.

The effects at different distances of various kinds of double discharge are also readily obtained from inspection of the two intersecting curves of fig. 3. For example, a discharge from a height of 2 km. to a height of 1 km., followed by an equal discharge from the lower level to earth, would produce at the surface of the ground two successive sudden changes of potential of the same or of opposite sign according as the distance exceeded or fell short of the above limit. Again if we consider a thunder-cloud of which the upper and lower parts are oppositely charged, and suppose that a discharge between the top of the cloud and the ground is followed by one between the ground and the bottom of the cloud, the two successive sudden changes of potential gradient would be of opposite sign, but their relative magnitude would depend on the distance of the place of observation from the discharges; at great distances the longer discharge, at small distances the shorter would produce the larger sudden change of potential gradient; while at some intermediate distance the two effects would be of equal magnitude. The various types of double discharge records shown in Plate 5 may perhaps be explained in this way; a given type of double discharge giving a considerable variety of effects on the trace according to its distance from the recording instrument.

If the effects of individual discharges could simultaneously be recorded at several suitable distributed stations, we should be able to learn much about the quantities of electricity which pass and about the initial and final distribution of charges. It is especially useful to have measurements of the change of field (1) at points at a considerable distance from a discharge, since the electric moment 2QH or $2Q(H_2-H_1)$ may at once be deduced, and also (2) for points nearly below the centres of the regions discharged, where, in the case of discharges to earth, F approximates to its maximum value $2Q/H^2$. Knowing both $2Q/H^2$ and 2QH we obtain both Q and H.

When a single station only is available we have to be content with attempting to learn something about the average lightning discharge by accumulating measurements of the effects produced by discharges at various distances.

VI. *Electric Moments of the Discharges.*

For each discharge recorded in fig. 2, FL^3, the product of the vertical electric force and the cube of the distance of the discharge, has been calculated. This product, when the quantities are expressed in electrostatic C.G.S. units, may be taken as giving a lower limit for the electric moment $2\Sigma qh = 2QH$, or $2Q(H_2 - H_1)$, becoming equal to it when L is large compared with H.

The mean value of FL^3 for the 78 positive flashes for which L could be determined in the 1917 records is $7{\cdot}3 \times 10^5$ in volts per metre × kilometres^3; for the 46 negative flashes the mean value obtained is identical with that found for the positive. We may take this value (the equivalent of $2{\cdot}4 \times 10^{16}$ E.S.U. × centimetres, or about 80 coulomb-km.) as a minimum estimate of the average electric moment of the lightning discharges.

The observations of 1917 give values of FL^3 ranging between 1/20 and 5 times the mean; but in more than half the discharges for which the necessary data are available FL^3 lies between one half and twice this mean value. Some of the eye observations made previously to 1917 lead to higher values, reaching in one case ten times the above mean.

In Table I. are given the mean values of FL^3 for positive and for negative discharges at distances (1) below 5 km., (2) between 5 and 10 km., and (3) exceeding 10 km. The number of observations used in getting the means is in each case inserted in brackets. FL^3 is given in volts per metre × kilometres^3 × 10^5.

TABLE I.

	Below 5 km.		5–10 km.		Above 10 km.	
	Positive.	Negative.	Positive.	Negative.	Positive.	Negative.
1917 only . .	4·5 (17)	3·7 (8)	6·0 (32)	8·2 (29)	11·8 (29)	9·5 (8)
1914 1915 1917 } . . .	3·7 (37)	3·7 (8)	7·4 (48)	7·8 (33)	14·6 (38)	15·5 (15)

The mean values of FL^3 are not appreciably different for positive and negative discharges.

For discharges at distances between 10 and 15 km., the mean value of FL^3 for the 24 positive discharges of 1917 is $10{\cdot}8 \times 10^5$; if the 5 discharges of previous years are included the mean is $11{\cdot}8 \times 10^5$ volts per metre × kilometres^3. Data for negative discharges between 10 and 15 km. are almost lacking.

These numbers leave little room for doubt as to the order of magnitude of the average electric moments of the discharges. Distances below 5 km. are too small in comparison with the probable lengths of the discharges for FL^3 to serve as a measure of the electric moment. We may assume that the value of FL^3 for a discharge at a distance of 10 km. or more approximates to its electric moment. The mean value of the electric moment for both positive and negative discharges may be taken as not differing much from 10^6 in volts per metre × kilometres^3 = 3×10^{16} E.S.U. × centimetres or about 100 coulomb-km.

Higher values for the mean electric moment are obtained, as is evident from Table I., if the data from discharges at greater distances than 15 km. are used. The records of discharges at great distances may possibly give disproportionately large values for the mean electric moment for two reasons: (1) because at these distances discharges of small electric moment are unrecorded on account of the small magnitude of the charges of potential gradient produced by them; and (2) because it is only at great distances that discharges, which do not reach the earth and which may be of great vertical length and have large electric moments, produce effects proportional to their moments. The sign of the effect of such discharges is in fact reversed at small distances, and the magnitude of the sudden change of potential gradient produced becomes more nearly proportional to the height of the lower end of the discharge than to its vertical length (fig. 3).

Some additions to the data of Table I. were furnished by the storm of June 17, 1917, in which the distance and frequency of the flashes were too great to admit of the distances of the individual discharges being estimated. There was in this case (see p. 80) independent evidence as to the approximate distance of the storm when the trace containing records of 95 positive and 40 negative discharges within 10 minutes was obtained. Assuming the distance of the discharges to have been 17 km. we obtain for the mean value of FL^3, in volts per metre × kilometres^3, $5{\cdot}8 \times 10^5$ for the 95 positive discharges and $3{\cdot}9 \times 10^5$ for the 40 negative, corresponding to electric moments 2QH of $1{\cdot}9 \times 10^{16}$ E.S.U. × centimetres = 63 coulomb-km. and $1{\cdot}3 \times 10^{16}$ E.S.U. × centimetres = 43 coulomb-km. respectively. The discharges were observed to be approximately vertical and to pass between the base of the cloud and the earth.

VII. *Quantity of Electricity Discharged in an Average Lightning Flash.*

When the electric moment of a discharge is known, the order of magnitude of the quantity of electricity which passes in the discharge may be roughly estimated. We may assume that the average vertical length of any ordinary discharge is likely to be between 1 and 5 km. Thus if the average electric moment 2QH is 100 coulomb-km., we may estimate the average quantity discharged in a flash as being between 10 and 50 coulombs.

We get some further information about the discharges by considering the way in which F varies with L (fig. 2). The charge which feeds a lightning flash is evidently not generally derived from a widely extended horizontal sheet, as is shown by the rapid falling off in F at comparatively short distances from the discharge.

The curve shown in fig. 2 represents the relation which would hold between F and L in the case of the discharge of 20 coulombs to earth from a point at a height of 2 km.; the charge may be considered to have been distributed symmetrically within a sphere around this point. The curve represents the mean of the observations fairly well, except in the case of discharges at great distances.

The average magnitude of the sudden change of field produced by lightning discharges at any distance may be roughly calculated by assuming that the average lightning flash consists of a discharge of 20 coulombs to earth from a height of 2 km.

The average change produced in the potential gradient by a discharge at a distance of 10 km. is, it will be noticed, of the order of 1000 volts per metre, and for moderate distances beyond this it probably falls off approximately according to the inverse cube law. (It should perhaps be pointed out that the change of field referred to here is merely the difference between the initial and final values, before and after the passage of a single discharge. At distant points the amplitude of the short period oscillations will greatly exceed the difference between the initial and final magnitudes of the field. Such oscillations—the ordinary "atmospherics" or "strays"—are of too short period to be recorded by the methods of this research).

Discharges may be expected to occur (1) between the ground and the lower part of a thunder-cloud; (2) between the upper and lower parts of the cloud; (3) between the upper part of the cloud and the ground; and (4) upwards from the top of the cloud. Great differences in the vertical lengths and in the electric moments of discharges are therefore to be expected, and the manner in which F varies with L in the different storms furnishes some evidence of such differences. When, as in the records of June 12, 1917, FL^3 varies little with the distance and is besides relatively small, one is tempted to conclude that the vertical length of the discharges was small, that, for example, they passed between the ground and the base of the cloud. When on the other hand, as on August 15, 1915, or August 15, 1917, FL^3 continues to increase with increasing distance and reaches very high values, great vertical lengths would appear to be indicated for the discharges. Possibly the discharges of greatest vertical length may be those between the top of a thunder-cloud and higher levels of the atmosphere.

It is unfortunate that no records were obtained of the effects of discharges from clouds immediately overhead; such observations of the maximum values of F would have given useful evidence bearing on the height from which the discharges came. A discharge of 20 coulombs from a height of 2 km. would cause at the ground a maximum change of potential gradient of nearly 100,000 volts per metre.

Comparatively few determinations appear to have been made of the dimensions of lightning flashes. A few are quoted by HANN,* the length of vertical flashes to earth generally ranging from 1 to 3 km. It is only rarely, in the photography of lightning, that the distance of the flash has been recorded, so that its length may be deduced. Fig. 4 is a reproduction of a photograph taken with this object in view and for which the necessary data are available; it is, moreover, of interest in other ways. It was taken on May 22, 1918, at about 22h. 45m., the camera pointing north. The interval between the lightning flash and the moment when the thunder began to be heard was 35 seconds, corresponding to a distance of 11·7 km. Two flashes are shown in the photograph, both passing between the cloud and the earth; they must have been nearly simultaneous, since the camera lens was covered as soon as a flash was observed. One discharge has initially passed upwards from the cloud and reached the ground by a curved path at a horizontal distance of nearly 4 km. from its starting point. The other has taken a nearly vertical course to the ground, its image is somewhat faint and ill-defined in the photograph: the discharge was probably within a heavy rain shower, a considerable thickness of which had to be traversed by the light on its way to the camera. The starting points of the two discharges in the cloud are comparatively close together, suggesting (as indeed does

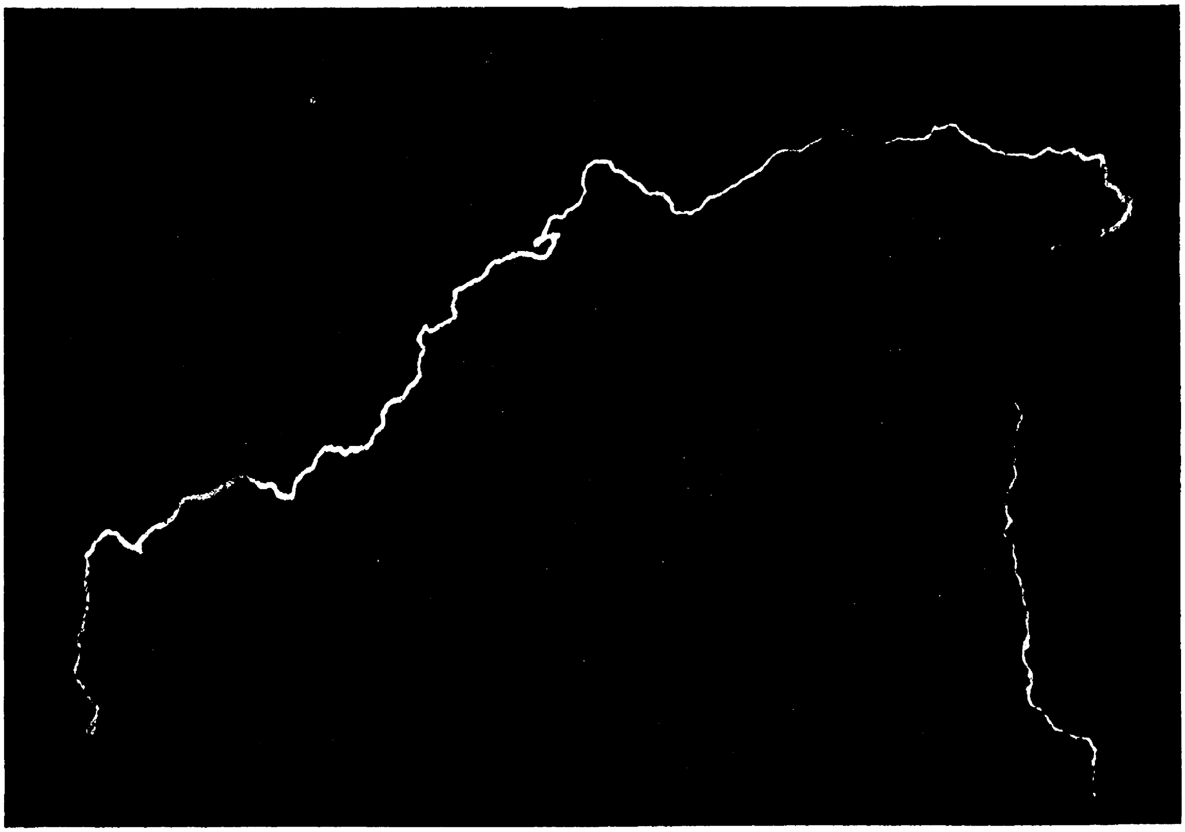


Fig. 4.

---

* HANN, 'Lehrbuch der Meteorologie,' p. 632, 1901.

the picture as a whole) that a charge of electricity had been concentrated in a comparatively small volume in the head of the cloud, and that the discharges took place approximately along lines of force.

The mean height of the upper ends of the two discharges—the height of the centre of the charged cloud-head according to this view—must have been just under 2 km., if its horizontal distance from the camera is taken as 11·7 km. The distance and height may in fact have been somewhat greater, since the track of the long flash may at some point of its course have been nearer the camera than the vertical flash, and the distance deduced from the interval between the lightning and thunder is that of the nearest point of the discharge.

VIII. *Electric Field of a Thunder-cloud.*

It is much more difficult to obtain direct information about the electric field of a thunder-cloud than about the sudden changes produced in the field by lightning discharges. The observed field may be the resultant of the fields of several thunder-clouds superimposed upon the normal electric field; while a single instantaneous change in the field will in general be due to the passage of one lightning flash, of which the approximate distance may frequently be determined. Nothing approaching a direct survey of the electric field of a thunder-cloud has yet been attempted: some general conclusions may be reached by a study of the photographic records of the potential gradient in thunderstorms.

It might perhaps naturally have been thought that the actual field due to a distant thunder-cloud would greatly exceed in magnitude the sudden changes due to the lightning discharges from it, each flash removing from the cloud only a small part of its whole charge. This is disproved by the observations; only when there has been, in addition to the more distant thunder-cloud, a heavy shower-cloud overhead or in the immediate neighbourhood of the place of observation, has the actual potential gradient greatly exceeded the instantaneous changes; the main part of the observed field has in all such cases obviously been due to the nearer cloud and not to the comparatively distant thunder-cloud which was in action at the time. The potential gradient due to a distant thunder-cloud has apparently never greatly exceeded in magnitude the sudden changes produced in the field by the lightning discharges from the cloud. Very frequently each discharge has approximately destroyed or even reversed the previously existing potential gradient, the field has then been rapidly regenerated, to be again nearly neutralised or reversed by the next discharge. The magnitude of the vertical electric force at the ground due to a thunder-cloud at a given distance is thus probably of the same order as has been found for the change produced by the average lightning discharge at the same distance.

Potential gradients exceeding 30,000 volts per metre (*i.e.*, $\frac{1}{100}$ of the sparking value) have not been recorded: it is doubtful, however, if any of the records were obtained when the centre of a storm was nearly overhead.

There can be little doubt that it is by the agency of precipitation that the separation of positive and negative charges in a thunder-cloud and consequent production of an electric field is effected, the larger raindrops or hailstones carrying down a charge of one sign while the charge of opposite sign is attached to small drops or cloud particles carried up in the ascending air stream. It is not proposed to discuss here how the large and small particles may acquire charges of opposite sign: whether for example the thunder-cloud is essentially a frictional electrical machine (disruption of drops, which SIMPSON* regards as the important factor, being included under this head) or an influence machine as ELSTER and GEITEL† contend.

It is obvious that any view that places the seat of electro-motive force of a thunderstorm within the thunder-cloud implies that the cloud is essentially bipolar, equal and opposite charges being in any given time transferred from within the cloud to its upper and lower portions. The actual charges residing at any moment in the positive and negative portions of the cloud will in general be quite unequal, since the conditions determining the rates of dissipation of the charges at the top and bottom of the cloud will be very different; an important part of the loss of charge from the lower part of the cloud is obviously the charge carried down to the ground in raindrops. The lower charge may indeed to a large extent reside on rain-drops falling from the cloud, and may thus extend all the way to the ground. Rain may not however reach the ground, or may lose a large part or the whole of its charge before reaching it by processes to be considered later.

Consider a cloud in which there is an upward stream of charged cloud particles or small drops and a downward stream of oppositely charged large drops; the total vertical electric current is the sum of the currents carried by the upward and downward streams. If the density of electrification of the two streams were the same and uniform throughout the greater part of the vertical thickness of the cloud, then the whole of this portion of the cloud would be electrically neutral. Above a certain level however the small drops alone will remain, and again it is only the large drops which fall below the lower margin of the cloud; equal and opposite charges will in this way be liberated in the upper and lower portions of the cloud. The assumption of uniform density of electrification in the two streams is of course an extreme and improbable one, and the concentration of the charges in the upper and lower parts of the cloud alone is not likely to be so complete as this supposition would imply; it serves, however, to indicate the possibility of the positive and negative charges of a cloud being separated by a considerable vertical thickness of electrically neutral cloud.

The factors which determine the rates of dissipation of the upper and lower charges and the magnitudes of the maximum charges are considered in a later section. The electric field at the ground due to a cloud of this kind will be the resultant of the fields of the upper and lower charges.

---

* SIMPSON, *loc. cit.*

† ELSTER and GEITEL, 'Wied. Ann.,' 25, p. 116, 1885; 'Physikal. Zeitschr.,' 14, p. 1287, 1913; GEITEL, 'Physikal. Zeitsch.,' 17, p. 455, 1916.

In the ordinary thunder-cloud or cumulo-nimbus cloud we are concerned with rapidly ascending air currents of comparatively small horizontal dimensions. The heads of such clouds generally reach to heights of several kilometres: according to WEGENER* the top of a thunder-cloud may reach almost to the upper limit of the troposphere (about 10 km.). The average height of the bases is probably about 1 km.

If we suppose a cumulo-nimbus cloud to have charges Q_2 and Q_1 of opposite sign in its upper and lower portions, we may, for the purpose of calculating its electric field at a distance, treat these charges as if they were concentrated at definite "poles" at heights H_2 and H_1. The effect of the charges induced on the surface of the ground is the same as if they were replaced by charges equal and opposite to Q_2 and Q_1 and at depths H_2 and H_1 below the surface. The problem is then the same as that of finding the magnetic field due to two bar magnets of lengths $2H_1$ and $2H_2$ and moments $2Q_1H_1$, $2Q_2H_2$, placed so that their centres coincide, the axes being vertical and their polarities opposed.

The vertical electric force due to the cloud at a point on the ground at a distance L from the axis is

$$F = \frac{2Q_2}{H_2^2\left(1+\frac{L^2}{H_2^2}\right)^{\frac{3}{2}}} - \frac{2Q_1}{H_1^2\left(1+\frac{L^2}{H_1^2}\right)^{\frac{3}{2}}}.$$

Immediately below the cloud, where $L = 0$, the second term (representing the effect of the lower charge) will be the greater unless the ratio of Q_2 to Q_1 exceeds H_2^2/H_1^2, and for distant points the first term (representing the vertical force due to the upper charge) will be the greater unless Q_1/Q_2 exceeds H_2/H_1. Thus the surface of the ground may generally be divided into two areas, an inner and outer, in which the electric field due to the cloud has opposite signs; in the central area the effect of the lower pole of the cloud predominates and determines the sign of the potential gradient and of the charge on the ground, while in the outer area the effect of the upper pole of the cloud is the greater.

The maximum intensity of the resultant field anywhere near the centre of the inner area will generally greatly exceed the maximum reached in the outer area. The curve III., fig. 3, represents the resultant potential gradient produced at the ground by equal and opposite charges of 20 coulombs at heights of 1 and 2 km. The inner area has a radius of approximately 2 km., the maximum potential gradient at the centre amounts to 270,000 volts per metre, while the maximum reached by the potential gradient of opposite sign in the outer area is less than 10,000 volts per metre. Greater differences in heights of the two poles are probable in actual thunderstorms, and the difference in the intensities of the electric fields in the inner and outer areas is likely to be even greater than in the example given. As represented in fig. 5 lines of force from the central area end on the lower charged portion of the cloud, those from the outer area on the upper charge, others again connect the upper and lower charges.

---

* WEGENER, 'Thermodynamik der Atmosphäre,' p. 210.

Thus far no account has been taken of the conducting layer in the higher levels of the atmosphere, to the existence of which the phenomena of terrestrial magnetism seem to point.

The normal potential gradient at the surface of the ground in clear weather is of the order of 100 volts per metre, falling off with increasing height and becoming negligible above 10 km.; thus the potential in the conducting layer over regions of fine weather is not likely to exceed a value of the order of 1,000,000 volts. If we assume, in accordance with modern theories of terrestrial magnetism,* that the conductivity of the upper atmosphere is high enough to prevent any large potential differences within it, then even above a thunderstorm the potential in the conducting layer may not greatly exceed 1,000,000 volts. The potential in the head of a thunder-cloud probably reaches values 1000 times as great.

One important effect of the conductivity of the upper atmosphere is to cause a portion of the lines of force from the head of the thunder-cloud to end in the conducting layer. The effect will be more marked than that which would be produced by a solid conducting sheet since ions of opposite sign to the charge on the head of the cloud will be dragged down out of the conducting layer to form an expansion of it extending downwards towards the thunder-cloud. The charge on these ions (which constitutes the induced charge on this protuberance from the conducting layer) will partially neutralise the electric field produced at the ground by the charge in the head of the cloud; in other words lines of force from the head of the cloud which would otherwise have ended on the ground are now diverted upwards into the conducting layer.†

The considerations brought forward in this section suggest that the electric field of a cumulo-nimbus cloud may be regarded as due to charges, generally unequal, in the upper and lower parts of the cloud (falling rain being included as part of the cloud) and to the charges induced by these on the ground and on the conducting layer of the upper atmosphere. Thus the lines of force of the cloud may be divided into four classes, connecting (*a*) the ground and the lower charge of the cloud, (*b*) the lower charge and the upper charge, (*c*) the upper charge and the

---

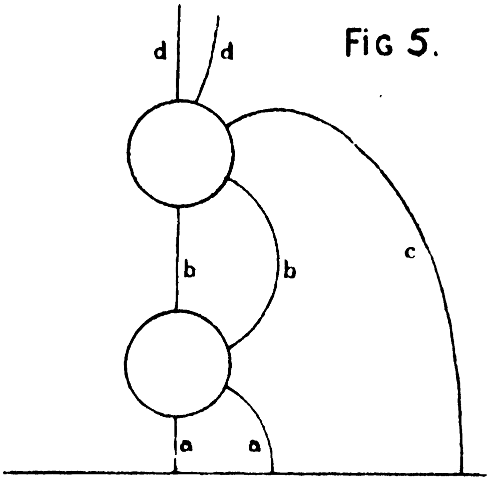
[*] SCHUSTER, 'Phil. Trans.,' A, vol. 208, p. 163, 1907; S. CHAPMAN, 'Phil. Trans.,' A, vol. 218, p. 1, 1919.

[†] In the absence of a previously existing conducting layer a thunderstorm would itself produce ionization in the upper atmosphere; this is readily seen to follow from the values found for the electric moments of lightning flashes.

ground, (*d*) the upper charge of the cloud and the conducting layer of the atmosphere.

If uniform stratiform conditions over a wide area be assumed, the conditions are simpler than in the case of the cumulo-nimbus cloud. The field at the ground below such a cloud, if the effects of the conducting layer be ignored, would be the difference between the fields due to the upper and lower charges, and its sign would be that of the field due to the larger charge. The effect of the conducting layer, as in the case of the cumulo-nimbus cloud, is to reduce the potential gradient produced at the ground by the upper charge of the cloud: firstly by the action of the opposing field of the charge induced on the conducting layer above the cloud, and secondly by the actual diminution of the cloud charge by the ionization current from the conducting layer.

FIG 5.



IX. *Conditions Determining Discharge.*

In order that a lightning discharge may begin, it is clear that the electric force must somewhere exceed the sparking limit, which amounts at the ordinary atmospheric pressure to about 100 electro-static units or 3,000,000 volts per metre; it is not necessary that the electric force along the whole length of the path of discharge should previously have approached the sparking limit. As LARMOR has pointed out,* if we suppose that an initial discharge occurs along a narrow line of length equal to the distance (possibly very small) over which the sparking value of the electric force was originally exceeded, and that this approximately equalises the potential along its path, there will be concentration of charge and intense local fields at the ends of this line; the discharge will thus be lengthened. The conditions are in fact momentarily much the same as if a conducting wire were placed along the path of this initial discharge. The maximum value of the electric force at the ends of the conducting track of the initial discharge will thus greatly exceed the critical value and will continue to do so as the track lengthens, so that the discharge may finally extend far beyond the boundary of the region in which the critical electric force was originally exceeded. Consider a stratiform cloud in which vertical separation of positive and negative electricity is taking place so that opposite charges are accumulating in the upper and lower portions of the cloud. Let us suppose that these charges remain approximately equal. There will be a vertical electric force within the cloud reaching a maximum in the central neutral zone of the cloud; the vertical electric force at the ground will be small and the conditions for discharge will be first reached within the cloud. Discharge will occur when the maximum vertical electric force within the cloud reaches its critical value; this amounts to about 30,000 volts per centimetre (= 100 electrostatic units) at a pressure of one atmosphere and is proportional to the pressure.

---

* Sir JOSEPH LARMOR and J. S. B. LARMOR, 'Roy. Soc. Proc.,' A, vol. 90, p. 312, 1914.

It is perhaps doubtful whether the vertical potential gradient within a cloud has necessarily to reach the above value of 3,000,000 volts per metre, in order that a discharge may begin, since an electric force amounting to one-third of this would be sufficient to bring the maximum electric force at the surface of a suspended drop, assumed spherical, to the above value. It must however be remembered that the critical value of the electric force at a curved surface of a conductor increases rapidly with the curvature and that only drops of the largest size will have any marked effect in assisting discharge.*

The discharge may extend considerably beyond the limits of the zone in which the vertical electric force originally reached the critical value. It is possible that it might extend even beyond the upper and lower boundaries of the cloud, for the ends of a linear vertical discharge would, as it lengthened, be continually penetrating into regions of a greater potential difference until they reached the limits of the charged portions of the cloud, so that the density of electrification and maximum electric force at the ends of the conducting track, and the consequent tendency to further lengthening of the discharge, would be increasing up to this point.

The end of an initial discharge which has penetrated into a region where there is little electric force to guide it will tend to branch or to expand into a brush. The potential may thus finally be approximately equalised throughout a considerable volume at each end of the discharge, the effective electric capacity of the expanded ends of the discharge and the original difference between the potentials of the regions thus connected will determine the quantity of electricity discharged by the complete flash.

If the lower charge of the stratiform cloud reaches nearly or quite to the ground

---

* The electric force at the surface of a conducting sphere 6 mm. in diameter has to reach about 260 E.S.U. (a value equal to nearly three times the sparking limit for a uniform field) in order that a spark may pass. (SCHUSTER, 'Phil. Mag.,' vol. 29, p. 182, 1890.) Drops of this size—which is little short of the maximum attained by rain drops—will only slightly assist discharge in an electric field in which they are suspended, and drops smaller than 3·5 mm. will not assist discharge at all.

(as will generally be the case when rain is falling), or if its charge is considerably smaller than the upper charge, then the initial discharge is likely to extend downwards till it reaches the ground; it will form then a conducting path for the main discharge, which may be regarded as reducing approximately to zero the whole discharge track and its ramifications.

If the potential gradient at the ground reaches a value amounting to any considerable fraction of that in the cloud, if for example, the upper and lower charges of a stratiform cloud are very unequal, then the critical value of the electric force is likely to be first reached, and the initial discharge to begin, at the surface of a projecting earth-connected conductor.

Let us next assume—and this perhaps represents more nearly the conditions which hold in an ordinary thunder-cloud—that the vertical separation of the centres of the charges is as great as the horizontal dimensions of the charged portions of the cloud.

Consider for example a charge to accumulate in the head of a cumulo-nimbus cloud until the conditions for the passage of a lightning discharge are reached. To get an idea of the order of magnitude of the quantities involved let us assume that the charge is distributed symmetrically about its centre within a sphere of radius R, the maximum electric force being at the boundary of the sphere. If the total charge of the sphere is Q, the radial electric force exerted at its surface by the charge is Q/R^2 and is there a maximum. A radial discharge will therefore begin at a point on the boundary of the sphere when F exceeds the critical value and will be continued inwards towards the centre and outwards approximately along a line of force. The charge of opposite sign in the lower part of the cloud will increase the electric force below and diminish it above the upper charge; the effect will however be small if the lower charge is small or if it is situated at a height small compared with that of the upper charge; in the latter case the effect of the lower charge is largely neutralised by the force due to its image. On the other hand on account of the diminished pressure at the greater height a smaller electric force is required to start a discharge from the upper than from the under side of the upper charged portion of the cloud.

Thus discharges may be expected to start not only downwards but also upwards and laterally from the charged head of a thunder-cloud. The path of discharge is likely to follow approximately a line of force which may belong to any of the classes of Section VIII. Discharges such as that of fig. 4, or even discharges upwards into a cloudless sky, such as have sometimes been observed, are not unlikely occurrences.

If an initial discharge from the charged head of a cloud reaches the ground, thus opening up a conducting path to earth, an approach to complete discharge is probable, so that the quantity of electricity which passes in the lightning flash may be taken as a measure of the charge which had accumulated in the head of the cloud.

A discharge originating in the region surrounding the lower pole of a cumulo-nimbus cloud is more likely to begin in the lower rather than the upper boundary of the charged region; since the electric force below will be increased, and that above diminished by the action of the induced charge, *i.e.*, virtually by the image of the lower charge. An extreme case is that in which the lower charge, carried largely by rain below the actual cloud, extends to the ground. Here the maximum value of the electric force would be at the surface of the ground; and, if the charge be assumed to be distributed uniformly throughout a region of which the vertical and horizontal dimensions are approximately the same, the maximum vertical electric force would not differ much from $2Q/R^2$, where R is the height of the centre of the lower charge. In this case the field will be locally intensified at the surfaces of projecting parts of earth-connected conductors, and discharges (not necessarily developing into lightning strokes) will occur from such points long before the electric force over flat ground reaches the sparking limit.

X. *Dimensions of the Regions Discharged by Lightning Flashes.*

It has been shown that the quantity of electricity which passes in an average lightning discharge—if the thunderstorms investigated may be taken as typical—is of the order of 20 coulombs. In this and the following sections, X. to XVII., are considered some of the consequences which follow if the quantity discharged by a lightning flash is taken as 20 coulombs.

Consider a thunder-cloud of the bipolar type and assume that a discharge takes place when the electric force at the boundary of either the upper or the lower charge reaches the sparking limit F_0. To get an idea of the order of magnitude of the effects, let us assume that the charge is contained within a sphere of radius R, at a distance from the ground and from other charged masses, and that it is distributed symmetrically in such a way that the maximum radial electric force is at the boundary. A discharge will occur when $Q/R^2 = F_0$. Thus, if $Q = 20$ coulombs $= 6 \times 10^{10}$ E.S.U. and $F_0 = 30{,}000$ volts per centimetre $= 100$ E.S.U. (its value at ground level) then $R = 250$ metres. If $F_0 = 50$, its value at a pressure of half an atmosphere, $R = 350$ metres. If an equal and opposite charge (the other cloud-charge or the image of the first in the ground) were similarly distributed within a sphere of the same radius in contact with the first, we should have at the moment of discharge $2Q/R^2 = F_0$; and the values found for R would be $\sqrt{2}$ times as great as in the case considered, *i.e.*, 350 and 500 metres respectively.

A similar result is obtained if, instead of assuming the charge to have been distributed in a sphere, we suppose the vertical thickness of the charged portion of the cloud to have been small compared with its horizontal dimensions. Consider for example the case in which there are frequent flashes between the earth and the base of the cloud. We may picture the charged rain escaping from the base of the cloud as forming a charged layer which increases in thickness at a rate equal to the downward velocity of the drops. The vertical electric force at the upper and lower boundaries of the charged layer, due to its charge, will amount to $2\pi\rho d$ where ρ is the charge per unit volume and d is the vertical thickness. On this will be superimposed the electric force due to the upper charge of the cloud and that due to the induced charge on the ground; the first of these will increase the electric force at the upper surface and diminish that at the lower surface, while the second will increase the electric force at the lower and diminish that at the upper surface of the charged layer. If we assume the electric force below the lower charge to be greater than above it—as may easily be the case if the vertical thickness of the cloud (of cumulo-nimbus type) is great in comparison with the height of the lower charge—its magnitude will be between $2\pi\rho d$ and $4\pi\rho d$. A flash will occur when the vertical electric force reaches the sparking limit, *i.e.*, about 100 in electrostatic measure. If we assume the boundary of the lower charge to be a circle of radius r, and the quantity discharged to be 20 coulombs $= 6 \times 10^{10}$ E.S.U., r is between 350 metres and 500 metres, these being the limits obtained by putting $F_0 = 2\pi\rho d$ and $F_0 = 4\pi\rho d$ respectively.

It has thus far been assumed that the horizontal dimensions of the charged portions of the cloud are less than the distance apart of their centres, and that the greater part of the whole upper or lower charge of the cloud is neutralised by each discharge. Let us now suppose that there has been a uniform stratiform distribution of charges over a wide area. Take as an example the cases in which the upper and lower charges of the cloud are equal, the other extreme case in which one of the charges is very small compared with the other is not essentially different—the charge on the ground taking the place of the second cloud-charge. There will be a discharge when $4\pi\sigma = F_0$, σ being the total charge in a vertical column of unit area extending throughout the whole thickness of either charged portion of the cloud. If 20 coulombs are discharged in a lightning flash, and the whole thickness of a limited area of the charged portion of the cloud is discharged by the flash, the area A, discharged is such that $AF_0/4\pi = 20 \times 3 \times 10^9$; if the area discharged be assumed circular, and F_0 be taken as 100, the radius of the area discharged must be approximately 500 metres.

XI. *Maximum Potential Attained before the Passage of a Lightning Flash.*

The potential at the surface of the sphere, considered in Section X., will immediately before discharge be approximately $Q/R = F_0R$; other terms being relatively small may be neglected in estimating the order of magnitude of the potential. The potential at the centre of the sphere will exceed that at the boundary, the excess lying between zero and F_0R—these being the values in the limiting cases (1) in which the radial electric force within the sphere is zero, the charge being confined to the boundary, and (2) in which the radial electric force within the sphere reaches everywhere the sparking limit. (The case of uniform distribution of the charge within the sphere is intermediate, the excess being $\frac{1}{2}Q/R$). The potential at the centre thus lies between $Q/R = F_0R = \sqrt{QF_0}$ and twice this value.

If $Q = 20$ coulombs $= 6 \times 10^{10}$ E.S.U. and $F_0 = 50$ E.S.U. the potential at the surface of the sphere before discharge must reach $1 \cdot 7 \times 10^6$ E.S.U. $= 5 \times 10^8$ volts.

We may take 10^9 volts as giving the order of magnitude of the potential reached in a thunder-cloud before the passage of a discharge of 20 coulombs.

The order of magnitude of the potential required to cause a discharge remains the same even if the spherical distribution of the charge is departed from: the horizontal dimensions might, for example, considerably exceed the vertical so long as they did not much exceed the height of the charge above the ground.

Suppose next that there is a stratiform distribution of charges over a wide area, so that the lines of force are vertical. The conditions of discharge have already been discussed in Section IX.

If we assume that the mean vertical electric force along the whole length of the line of discharge initially approached the value F_0 (= about 3×10^6 volts per metre) and that this length is 2 km., the potential difference between the levels connected by the discharge must have been about 6×10^9 volts. But, as was pointed out in Section IX., the discharge may extend much beyond the regions in which the vertical electric force had originally attained the sparking limit F_0; the discharge might, for example, extend from the region of the upper charge of the cloud to the ground, although the electric field did not originally extend to the ground. The potential difference required to produce a vertical lightning flash 2 km. long from a cloud of this type may thus be considerably less than 6×10^9 volts, but it is not likely to be so small as 10^9 volts.

XII. *Mean Density of the Charge in a Thunder-cloud immediately before Discharge.*

If we assume that a charge of 20 coulombs is concentrated within a sphere 500 metres in radius, the charge per cubic metre is about 120 E.S.U.

In the case of a stratiform distribution of charges we have immediately before discharge $4\pi\sigma = F_0$ (Section X.). If uniform density ρ be assumed for the charge throughout a layer of thickness d, then $\rho d = F_0/4\pi$ = about 8 E.S.U. If d be taken as equal to 1 km., $\rho = 8 \times 10^{-5}$ E.S.U. per cubic centimetre (= 80 E.S.U. per cubic metre). Concentration of the charge within a smaller thickness is probable, with a corresponding increase in the density of the charge.

The mean density of the charges in thunder-clouds is thus likely to reach values of the order of 100 E.S.U. per cubic metre.

XIII. *Charge Associated with 1 c.c. of Water.*

If the amount of water in the charged portion of a thunder-cloud were no greater than in ordinary clouds (about 4 gm. per cubic metre), the average charge per gramme of water would be about 25 E.S.U.; the force exerted on each gramme of water by the electric field where it approached the sparking limit, 100 in E.S. measure, would amount to 2500 dynes, *i.e.*, to more than twice its weight. As pointed out by SIMPSON,[*] 10 E.S.U. is the largest charge per cubic centimetre of water consistent with its falling in an opposing electric field of 100 E.S.U. (on one occasion rain actually was found by him to carry a charge exceeding 10 E.S.U. per cubic centimetre). In the same paper SIMPSON draws attention to the very considerable accumulation of water that must occur in thunder-clouds through lagging of the larger drops behind the uprushing air. Thus the charge per cubic centimetre of water does not necessarily reach the above high values: and indeed the electric force opposing the fall of the large drops associated with the lower pole of the cloud cannot, as a rule, exceed their weight, since it is by the fall of these drops that the field is maintained. But there will be less concentration of water on the smaller drops associated with the upper charge, and densities exceeding 10 E.S.U. per cubic centimetre in the upper part of the cloud are not unlikely.

The drops in the head of a thunder-cloud may thus in virtue of their mutual repulsion have radial velocities which near the boundary may be comparable with the terminal velocity which the drops would acquire if falling freely through the air. Drops of 10^{-3} cm. in radius would have a maximum radial velocity of a few centimetres per second: if the radius were as large as 5×10^{-3}, the charge per cubic centimetre of water remaining the same, the radial velocity would be of the order of 1 metre per second. The characteristic bulging form of the heads of a developing cumulo-nimbus cloud may possibly be partly due to mutual repulsion of the charged droplets.

XIV. *Disruption of Drops by the Electric Field.*

It was shown by Lord RAYLEIGH[†] that a charged spherical drop must become unstable if Q^2 exceeds $16\pi a^3 T$, where Q is the charge, a the radius of the drop and T the surface tension. If the charge per cubic centimetre of the water in the cloud is ρ and is equally distributed among the drops, so that $Q = \frac{4}{3}\pi a^3 \rho$, then the spherical form will be stable so long as $\rho^2 a^3$ does not exceed $9T/\pi$, *i.e.*, about 225 in the case of water drops. The limit fixed in this way for the maximum charge per cubic centimetre of water, even for rain-drops as large as $\frac{1}{3}$ cm. in radius (for which it amounts to more than 70 E.S.U.), is too high to be of importance in the thunder-cloud problem.

Of much greater importance is the effect, upon the stability of the drops, of the electric field in which they are suspended, or in other words of the induced charges on the two halves of each drop.

If it is as a result of the electric force within or at the boundary of a cloud that a lightning flash occurs, then it becomes an interesting question whether under certain conditions disruption of the drops may not occur before the conditions for discharge are reached. ZELENY,* who has made a very interesting series of investigations on the stability of electrified liquid surfaces, found that in air at atmospheric pressure the potential required to cause a discharge from the surface of a drop of water at the end of a capillary tube exceeds, though only by a few per cent., that required to produce instability and disruption of the drops. He points out that it would follow from his experiments that a discharge of minute electrified drops, constituting an upward shower, would take place from the edges of the wet leaves of a tree in a thunderstorm, before the electric force at the surface of the tree reached the sparking limits.

---

[*] SIMPSON, *loc. cit.*

[†] RAYLEIGH, 'Phil. Mag.,' vol. 14, p. 184, 1882.

It seems not unlikely that under certain circumstances a similar process may occur in a cloud, droplets suffering disruption where the field approaches the sparking limit. Consider a developing cumulus head in which a charge is accumulating, and suppose that the radial electric force near the edge of the cloud-head reaches the value required to cause disruption of the drops before it reaches the sparking limit. The induced charges on the two halves of the drop will then be separated and will tend to travel in opposite directions along a line of force.

The magnitude of the induced charge on each half of a spherical drop of radius a in a field F is $3\pi a^2\mathrm{F}/4\pi = \frac{3}{4}\mathrm{F}a^2$, and when F approaches the sparking value this will generally greatly exceed any resultant charge of the drop. The charge per cubic centimetre of water for each half of the drop $= \frac{3}{4}\mathrm{F}_0 a^2/\frac{2}{3}\pi a^3 = 9\mathrm{F}_0/8\pi a$; if $\mathrm{F}_0 = 100$ and $a = 1$ mm. the charge per cubic centimetre of water for each half of the drop is 360 E.S.U. Thus if the original drop of 1 mm. in radius were divided into two oppositely charged halves, the force acting on each of the new drops would in a field of 100 E.S.U. amount to 36 times its weight.

If the drop is drawn out by the action of the field into an ellipsoidal or cylindrical form before disruption, the induced charges will be considerably greater. Separation of the charges by division of the drop will thus give rise to oppositely charged portions each having a charge much greater than that of the original drop. The two portions will tend to travel in opposite directions along the line of force with velocities greatly exceeding the original radial velocity of the drop from which they were derived.

The outward moving products of disruption of the drops in the head of a cumulonimbus cloud may possibly constitute "false cirrus." These particles are more likely to freeze than those constituting the original head of the cloud; (1) because the stretching of the water drop into a filament itself causes cooling; (2) the conversion of a water filament into an ice crystal is not accompanied by a large increase of surface, and one of the main obstacles in the way of the freezing of small drops is thus removed; and (3) the charged particles are still further cooled through being driven by the action of the field into the colder and drier air outside the original cloud.

Ice needles formed under the above conditions would not only be charged but also electrically polarised (the induced charges of the original conducting filament remaining when the filament freezes), and will thus tend to remain with their long axes parallel to the direction of the electric force. A study of the optical phenomena of "false cirrus" would be of interest in this connexion, as would also an experimental investigation of the effects of an electric field on super-cooled drops.

---

* ZELENY, 'Proc. Cambridge Phil. Soc.,' vol. 18, p. 71, 1914.

XV. *Pressure Within a Charged Portion of a Cloud.*

The pressure within a charged cloud—like that within a charged soap bubble—must be less than the pressure outside. If the whole charge be supposed to lie near the surface of a sphere the analogy with the soap bubble is complete, and this case may be considered in finding the order of magnitude of the effect. The reduction of pressure within the cloud by the charge is $2\pi\sigma^2 = F^2/8\pi$ where σ is the charge per unit area of the surface of the sphere and F is the radial electric force immediately outside. Just before the passage of a discharge $F = F_0$ = about 100 in electrostatic measure, so that $F^2/8\pi$ is about 400 dynes per square centimetre, *i.e.*, about $\frac{1}{2500}$ of an atmosphere.

If we consider the charge to be distributed uniformly in a horizontal layer of thickness which is small compared with its horizontal dimensions, the diminution of pressure midway between the top and bottom of the charged layer, due to mutual repulsion of the charged drops, is again $F_0^2/8\pi$ dynes per square centimetre.

XVI. *Thunder Resulting from Sudden Contraction due to Loss of Charge.*

Thunder is generally regarded as entirely due to the sudden expansion of the air along the track of a lightning flash. It is evident however that the sudden contraction of a large volume of air (the contraction corresponding to an increase of pressure of some tenths of a millimetre of mercury) must furnish a by no means negligible contribution to the thunder which follows the discharge.

XVII. *Energy Dissipated in Lightning Discharges.*

If we take the estimates arrived at above ($V = 10^9$ volts, $Q = 20$ coulombs) for the order of magnitude of the potential in the charged portions of a thunder-cloud immediately before the passage of a flash, and of the quantity discharged in the flash, we obtain for the order of magnitude of the energy dissipated in an average discharge, $\frac{1}{2}QV = 10^{10}$ joules $= 10^{17}$ ergs.

We may also arrive at an upper limit for the energy if we assume that the distribution of the charges is stratiform and that the vertical electric force is uniform and equal to F_0 throughout the height H through which the discharge extends. From the value found for the average electric moment, 2QH, since V must now be equal to F_0H, we have for the energy dissipated, $\frac{1}{2}QV = \frac{1}{2}QF_0H$, about 10^{11} joules.

The rate at which electrical energy would be going to waste in a storm in which one such discharge occurred in every 10 seconds would amount to 10^{16} ergs per second or 1,000,000 kilowatts. It is of interest to compare this with the total power which would be available if it were possible to catch the rainfall of a thunder-shower before it fell and utilise the water power thus stored. The rate of rainfall in a severe thunderstorm may reach values approaching 10 cm. per hour. The water power available if it were possible to catch the rain at a height of 1 km. would amount to 3×10^{15} ergs per sq. kilometre per second. Thus a rainfall of the above amount over an area of about 3 sq. km. if intercepted at a height of 1 km. would furnish sufficient power to produce the required electrical energy. The total power available for the production of lightning flashes may obviously greatly exceed the above estimate based on the rainfall.

XVIII. *Interpretation of "Recovery" Curves.*

In a typical record of the changes of the vertical electric force due to a distant thunderstorm each vertical portion of the trace—representing the sudden change produced by a discharge—is followed by a characteristic "recovery" curve. This may be interpreted as follows:—The charge in the head or base of the thunder-cloud—or in both—is suddenly destroyed by the passage of a lightning flash. The field at once begins to be re-established at a rate represented by the initial steepness of the curve immediately after the discharge. But as the charge increases, its field tends to diminish the rate of increase of the charge in two ways: (1) by hindering the separation of oppositely charged rain-drops and cloud particles; and (2) by producing an ionization current which tends to neutralise the charge and increases with the increasing intensity of the field. Unless the field previously reaches the sparking limit, a steady condition will finally be approached when the two opposing processes, which tend respectively to increase and diminish the field, balance one another.

The initial rate of increase of the field immediately after the passage of a distant discharge is thus an important quantity. It is proportional to the rate at which a charge destroyed by the flash is regenerated by the action of the thunder-cloud, *i.e.*, it is proportional to the vertical electric current which is carried through the thunder-cloud by the convection of charged masses. If the distances and height of the charge destroyed are known, the vertical electric current may at once be deduced from the initial rate of increase of the field. If this information is not available the ratio of the current to the quantity which passed in the previous discharge can always be obtained from the record.

The value of $T = F\Big/\frac{dF}{dt}$, where F is the instantaneous change recorded and dF/dt the initial rate of recovery immediately after the discharge, has been deduced from the recovery curves in the case of 34 discharges. We may regard T as the time which would have been required to re-charge the cloud to the sparking limit had there been no neutralising process due to the action of the electric field of the cloud. The values of T vary between 1·5 seconds and 30 seconds, the mean of 64 measurements giving 6·9 seconds; in more than half the cases examined T lies between 4 and 10 seconds. These times are generally only a small fraction of the actual intervals between the flashes: on June 17, however, in a record (Plate 3, fig. 5) showing more than 100 flashes in 10 minutes—so that the average interval between the flashes was less than 6 seconds—the average value of T exceeded half this interval.

Some of the recovery curves, as, for example, that of June 12, shown in Plate 4, fig. 11, approximate very closely to the exponential form, so that the charge which has been regenerated when a time t has elapsed after the discharge may be represented by $Q = Q_0(1-e^{-\lambda t})$. Such a curve suggests that the charge of the thunder-cloud is being regenerated at a constant rate, and that it is at the same time being dissipated at a rate which is at any moment proportional to the charge. It might also however be interpreted as representing the regeneration of the charge by a constant E.M.F. in the cloud, the current through the cloud being proportional to the difference between this E.M.F. and the opposing potential difference produced by the charges separated; there would be no current when the charges reached a steady value. If dissipation of the accumulated charges is taken into account the recovery curve still remains of the same type; if the dissipation is large, or, in other words, if the internal resistance of the thunder-cloud, regarded as a generator of constant E.M.F., is large compared with that of the external circuit, the current through the cloud is constant, and we have again the case first considered.

The rate of regeneration of charge per second, in other words the vertical current through the cloud, immediately after a discharge varies between $\frac{2}{3}$ and $\frac{1}{30}$ of the charge removed by the flash, the mean being about $\frac{1}{7}$. If we assume a discharge to convey a quantity of the order of 20 coulombs, the mean current through the cloud, immediately after a discharge, is of the order of 3 ampères.

It is not at all impossible that this is also the order of magnitude of the vertical current through a thunder-cloud at other times than immediately after a lightning discharge, and even when an approximately steady condition of the field has been reached. Consider, for example, the charge in the head of a thunder-cloud which reaches to a great height. The conductivity of the atmosphere has been found by GERDIEN and by WIEGAND* to increase rapidly with the height, the former having found at 6 km. a conductivity more than 20 times as great, and the latter at 8865 metres a conductivity about 40 times as great as the normal conductivity near the ground. A charged body suspended in the atmosphere under the conditions found by **Wiegand** at 8865 metres would lose about $\frac{1}{15}$ of its charge per second. Thus a charge of 20 coulombs in the head of a thunder-cloud at this height should lose more than 1 coulomb per second: to keep the charge constant the vertical current through the cloud would have to exceed 1 ampère. The presence of such a large charge would, it is true, not leave the conductivity of the surrounding atmosphere unaltered: it would tend to increase it by dragging down ions from upper layers of still greater conductivity.

---

* WIEGAND, 'Deutsch. Physik. Gesellschaft,' February 29, 1914.

XIX. *Electrical Currents Maintained in the Atmosphere by Thunder-clouds and Shower-clouds.*

Consider a cumulo-nimbus cloud of the type imagined in Section VIII. containing upper and lower charges—the latter being partly or, it may be, mainly carried by the rain below the cloud. Such a cloud may be regarded as an electric generator—whether essentially of the frictional type or of the influence machine type need not at present be discussed—capable of maintaining a potential difference between its poles of the order of 10^9 volts.

As pointed out in Section VIII. the potentials in the conducting layer of the upper atmosphere is likely to be insignificant in comparison with that in the head of a thunder-cloud, and the potential difference between them may thus be of the order of 10^9 volts.

There will be a flow of electricity along the lines of force belonging to the various groups enumerated in Section VIII. and indicated in fig. 5. The upper pole will continually be losing charge by currents flowing (1) to the lower pole; (2) to the earth's surface (this portion of the current reaching the outer zone (Section VIII.) where the potential gradient is unlikely to reach high values); and (3) to the upper atmosphere.

Unless the field in the shower-cloud approaches very near to the sparking limit, the conductivity within the cloud is likely to be small, since any ions liberated soon lose their mobility by becoming attached to cloud particles. The electrical resistance of the atmosphere between the upper pole of the cloud and the conducting layer of the upper atmosphere will be much less than that between the upper pole and the earth's surface; for the free ions will be dragged out of the conducting layer, and their mobility throughout the greater part of their course will greatly exceed that of the ions in the lower layers of the atmosphere. A large part of the current from the upper pole must thus go to the upper atmosphere.

Consider now the lower oppositely charged pole of the cloud. Part of the charge is continually being neutralised by the direct return current between the poles, but this, as has already been pointed out, is likely to be small. The greater part of the charge lost by the lower pole will reach the ground. If no rain reaches the ground the loss of charge will be due to ions moving under the action of the electric field of the cloud. If the normal rate of production of ions in the air below the cloud had alone to be taken into account, the current would be small; but we have to add the ions supplied by evaporation of charged drops falling from the cloud and those (of opposite sign) due to point discharges from earth-connected conductors, such as the leaves of trees or even the tips of blades of grass, under the action of the intense electric field of the central area below the cloud. If the rain reaches the ground the former of these sources of ionization is absent, but there is a further source of ionization in the splashing of the rain on the ground. In addition to the ionization current we have also the convection current carried to the ground by charged rain-drops. The total current between the lower pole of the cloud and the ground now consists of the convection current carried by the falling charged drops and the conduction current carried mainly by the upward stream of ions set free by point discharges and splashing at the surface of the ground. The ratio of the convection current to the conduction current will be less near the ground than higher up, since the falling drops will lose more and more of their charge as they penetrate farther into the stream of upward moving oppositely charged ions; these again as they are carried upwards by the electric field are continually diminished in number by union with the drops. The greater the supply of ions from the ground the smaller will be the charge retained by the drops; if the current carried by the upward stream of ions is sufficient, the drops may lose the whole of their charge or even have it reversed before they reach the ground. The charge carried to the ground by rain-drops is thus by no means necessarily a true measure of the vertical current in a shower: nor does the sign of the charge carried by the drops when they reach the ground necessarily indicate the sign of the current between the ground and the base of the cloud.

Thus a large part of the current from the upper pole of a cumulo-nimbus cloud is likely to reach the conducting layers of the upper atmosphere, while that from the oppositely charged lower pole goes mainly to earth. A current is thus maintained from the earth through the cloud to the upper atmosphere or in the reverse direction according to the sign of the polarity of the cloud.

Discharges between the ground and the lower pole of the cloud and between the upper pole and higher portions of the atmosphere contribute to the total current between the ground and the upper atmosphere; discharges between the two poles or between the upper pole and the ground diminish the electric field which maintains the vertical current without contributing anything to the current.

XX. *Differences Between the Electrical Effects of Shower-clouds of Positive and Negative Polarity.*

We may define the polarity of a shower-cloud as being positive when the upper charge is positive, negative when the upper charge is negative, the current through it being upward in the former case, downward in the latter.

It was first proved by SIMPSON,[*] and has been confirmed by many observers, that rain on reaching the earth's surface is much more often positively than negatively charged. This, as we have seen, does not necessarily imply that shower-clouds are always or even prevailingly of negative polarity. It is therefore of interest to consider some of the differences to be expected between the electrical effects of clouds of positive and of negative polarity.

Recent experiments have shown[†] that the carrier of negative electricity in hydrogen, helium and nitrogen even at atmospheric pressure is the free electron, and that its mobility is some hundreds of times that of the carrier of positive electricity, the positive ion. In ordinary atmospheric air, as the pressure is reduced, the average mobility of the carriers of negative electricity increases relatively to that of the positive ions; quite an appreciable proportion of the negative carriers, consisting, according to WELLISCH,[‡] of free electrons when the pressure is reduced to 8 cm. of mercury, the proportion increasing rapidly as the pressure is further reduced.

Thus, while the carriers of positive electricity dragged out of the conducting upper atmosphere by a cloud of negative polarity consist of ordinary ions, the negative carriers dragged down by a cloud of positive polarity are originally to a large extent free electrons, and a considerable proportion are likely to remain in this condition till quite moderate elevations are reached. The conductivity of the air between a shower-cloud and the upper atmosphere will thus be considerably greater if the cloud is of positive than if it is of negative polarity.

Let us compare two shower-clouds which differ only in the sign of their polarity and consider the effect of the greater conductivity of the atmosphere above the cloud of positive polarity. Let us suppose that the two clouds act as generators capable of maintaining equal potential differences between their poles. Let V_2, $-V_1$ be the potentials of the upper and lower poles of the cloud of positive polarity, and $-V_2^1$, $+V_1^1$ the potentials of the upper and lower poles of the cloud of negative polarity, let $V_2-V_1 = V_1^1-V_2^1$. Then the current from the ground to the upper atmosphere maintained by the cloud of positive polarity will be greater than that from the upper atmosphere to the ground maintained by the cloud of negative polarity, since the total resistance of the circuit is less in the former case.

The ratio V_2/V_1 is less than V_2^1/V_1^1, the upper and lower potentials being proportional to the resistance of the portions of the circuit above the upper and below the lower pole respectively.[§] Thus V_2^1 is greater than V_2 and V_1 is greater than V_1^1; in other words the potential (and charge) of both the upper and the lower pole is greater when negative than when positive.

---

[*] SIMPSON, *loc. cit.*

[†] FRANCK and POHL, 'Verhandl. Deutsch. Physik. Gesellschaft,' 9, p. 69, 1907.

[‡] WELLISCH, 'Phil. Mag.,' vol. 34, p. 33, 1917.

[§] The potential of the conducting layers of the upper atmosphere is assumed to remain small in comparison with the E.M.F. of the thunder-cloud.

The potential gradient (negative) in the central area below the cloud of positive polarity will be greater than the positive potential gradient in the corresponding area below the cloud of negative polarity, the central positively charged area below the cloud of positive polarity being also larger than the negatively charged area below the cloud of negative polarity. Again, the positive potential gradient at the ground in the outer zone will be less (on account of the smaller charge on the upper pole) in the case of the cloud of positive polarity than the negative potential gradient in the corresponding region due to the cloud of negative polarity. Thus in each area negative potential gradients tend to be greater than positive.

The electric field in the central area below the lower pole being stronger in the case of the cloud of positive polarity, the current carried by the stream of positive ions from the ground will be increased, and therefore also the tendency to neutralisation or reversal of the negative charge on the falling rain-drops.

If lightning discharges occur, they are more likely to pass between the ground and either the upper or the lower pole if this is negative than if it is positive, since the charge of the pole is greater when negative. Thus discharges carrying positive electricity from the earth to the atmosphere will be more frequent than negative discharges. Discharges will tend to occur especially between the ground and the upper, negative, poles of clouds of negative polarity and the lower, negative, poles of clouds of positive polarity. In the latter case the discharges are an additional source of loss or reversal of the negative charge on falling rain-drops.

Essentially similar results are reached if, instead of assuming the same potential difference to be maintained between the poles, whether the clouds are of positive or of negative polarity, we assume that the same vertical current is maintained in both cases.

Thus, if we assume that shower-clouds may have polarity of either sign, the differences in the mobilities of the positive and negative carriers of electricity in the higher portions of the atmosphere will account for the preponderance in showers: (1) of negative potential gradients; (2) of upward or positive lightning discharges; and (3) of positively charged rain. It also affords (4) a possible explanation of the normal positive potential gradient of fine-weather regions.

XXI. *The Normal Potential Gradient and Air-earth Current of Fine Weather.*

A thunder-cloud or shower-cloud is the seat of an electromotive force which must cause a current to flow through the cloud between the earth's surface and the upper atmosphere. In the case of thunder-clouds the records of the changes produced in the electric field by the passage of lightning flashes give us means of forming some estimate of the magnitude of such currents, and it would appear from them that the current through a few square kilometres of the surface of the ground below a thunder-cloud may amount to some ampères. In shower-clouds in which the potentials fall short of what are required to produce lightning discharges, there is no reason to suppose that the vertical currents are of an altogether different order of magnitude. If any considerable proportion of shower-clouds are of positive polarity the upper atmosphere will receive an excess of positive electricity which may possibly be sufficient to maintain the positive potential of the conducting layers and to supply the normal downward current of the fine-weather regions. The total current which must be supplied for this purpose is, as SIMPSON* has pointed out, of the order of 1000 ampères for the whole earth.

It is not necessary to suppose that only isolated clouds of the cumulo-nimbus type contribute to the current between the ground and the upper atmosphere. If we consider a cloud from which heavy rain is falling and assume that the conditions are uniform over a large area, the case is in fact somewhat simpler than that of the cumulo-nimbus cloud; the general results are the same.

We may suppose that a steady condition is reached in which the vertical electric field within the cloud (and thus the potential difference between its upper and lower surfaces) has a value which depends on the rate of rainfall and other factors; it is assumed to be independent of the sign of the polarity. Even if this potential difference is only $\frac{1}{10}$ or $\frac{1}{100}$ of that reached in thunder-clouds the effects may be important: the E.M.F. which tends to drive a current between the ground and the upper atmosphere is still from 10 to 100 times the normal potential of the upper atmosphere above fine-weather regions.

The difference between the mobilities of the positive and negative carriers dragged out of the conducting upper atmosphere will again cause clouds of positive polarity to differ from those of negative polarity in (1) the greater magnitude of the vertical current (positive for the cloud of positive polarity); (2) the smaller magnitude of the potential (positive) at the upper surface and greater magnitude of the potential (negative) at the lower surface of the cloud; and thus (3) the greater intensity of the potential gradient (negative for the cloud of positive polarity) below the cloud, this again tending to cause a larger part of the vertical current below the cloud to be carried by ions liberated at the ground and thus to produce a more complete discharge of the (negatively-charged) rain.

XXII. *Influence of the Nature of the Earth's Surface below a Thunder-cloud or Shower-cloud.*

The dissipation of the lower charge of a thunder-cloud or other rain-cloud by the upward stream of ions liberated by point discharges or by splashing at the earth's surface must depend largely on the nature of that surface, on whether for example it consists of desert, snowfield, grassland, forest, lake or sea; and again the effect of the nature of the covering of the earth's surface may depend on the sign of the electric field.

---

* SIMPSON, 'Nature,' December 12, 1912.

Point discharges will occur most frequently and give rise to the largest currents over forests and lands covered with vegetation; also on mountain summits and ridges, owing to the increased intensity of the electric field through proximity to the charged cloud. Ionization by splashing of rain on the ground and the relative number of positive and negative ions liberated thereby is likely to be very different over the various surfaces. Of special interest is the question of the amount and nature of the ionization at the surface of the ocean under heavy rainfall.

Over an area in which the surface ionization was large we should expect an increased vertical current, a diminution of the charge carried to the ground by rain, a diminution in the intensity of the electric field of the cloud, and in consequence a diminution in the tendency for lightning discharges to occur.

The holding up of charged rain-drops by the electric field and the diminution of the field by the dissipation of cloud charges by forests and other sources of surface ionization are perhaps not negligible factors in the local distribution of rainfall. Mr. L. F. RICHARDSON,* describing some of the phenomena observed during the passage of a line squall in France, on September 6, 1917, remarks "the cloud was noticeably darker over the Forest of Argonne than over the grasslands of Champagne."

XXIII. *Secondary Thunder-clouds.*

It has thus far been assumed that the source of E.M.F. is within the cloud in which the lightning discharges and other electrical effects are manifested. It is easy however to imagine conditions in which a cumulo-nimbus cloud, which acts as electric generator, may supply electrical energy to quiescent clouds and produce in them intense electrical fields and even lightning discharges.

Consider for example a horizontal stratiform cloud which intersects lines of force connecting the poles of a cumulo-nimbus cloud; the stratiform cloud might be a lateral extension of the shower-cloud. The electrical conductivity within the stratiform cloud will, through capture of the ions by cloud particles, be very small compared with that of the free air above or below the cloud. The current from the poles of the primary thunder-cloud will cause an accumulation of charges of opposite sign at the upper and lower surfaces of the stratiform cloud. This accumulation of charge will continue—unless the field within the cloud previously reaches the sparking limit—until a steady condition is reached, when the vertical field within the cloud is sufficient to maintain a current equal to that which enters its upper and lower surfaces. The potential difference finally existing between the upper and lower surfaces of the stratiform cloud might amount to a considerable fraction of that between the top and bottom of the shower-cloud, the ratio being that of the resistance of that portion of a tube of flow which lies within the cloud to the resistance

---

* RICHARDSON, 'Quart. Journ. Roy. Meteor. Soc.,' 45, p. 112, 1919.

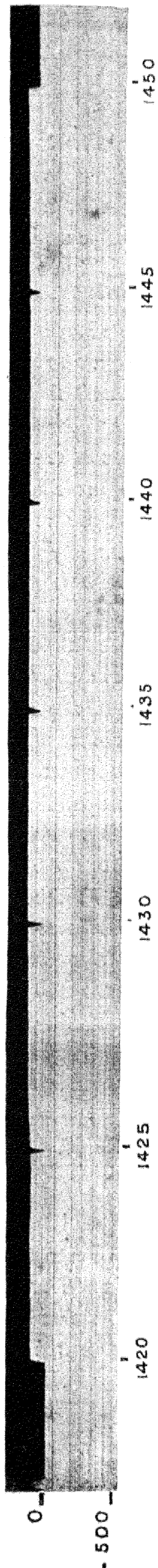


Fig. 1, May 23, 1917.

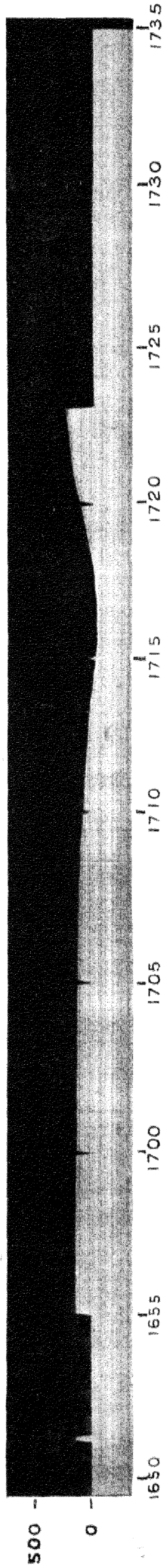


Fig. 2, May 12, 1917.

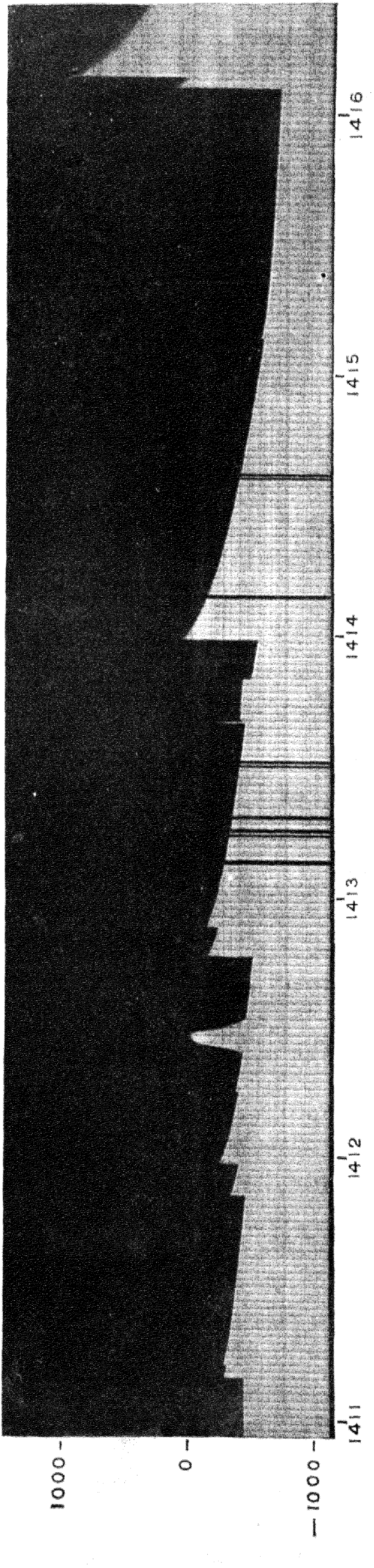


Fig. 3, June 13, 1917.

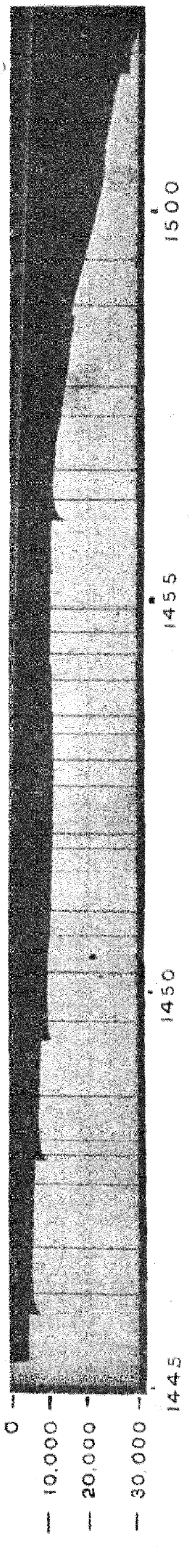


Fig. 4, August 9, 1917.

Fig. 5, June 17, 1917

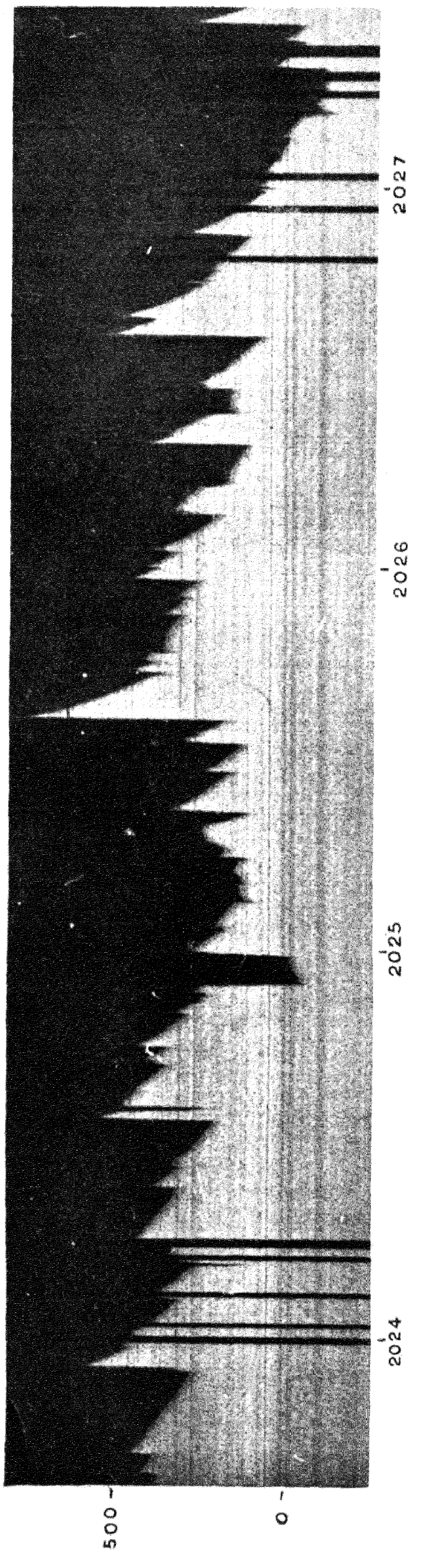


Fig. 6, June 16, 1917.

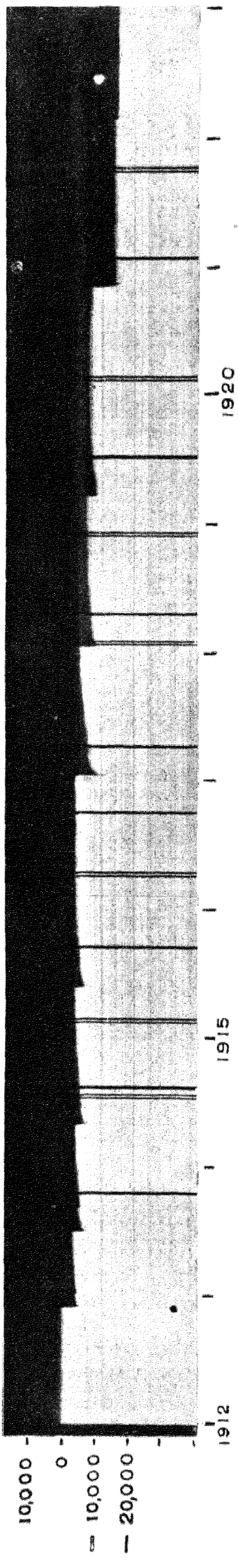


Fig. 8, May 29, 1917.

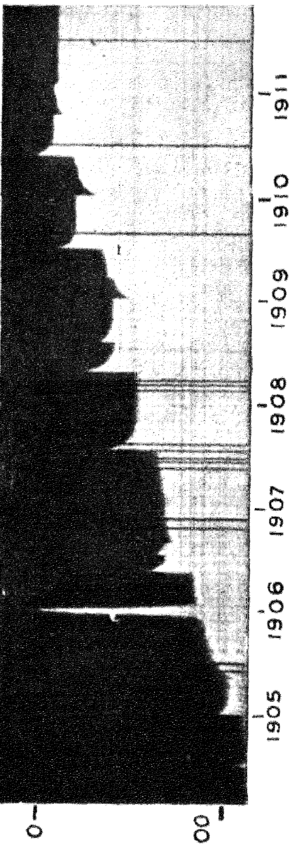
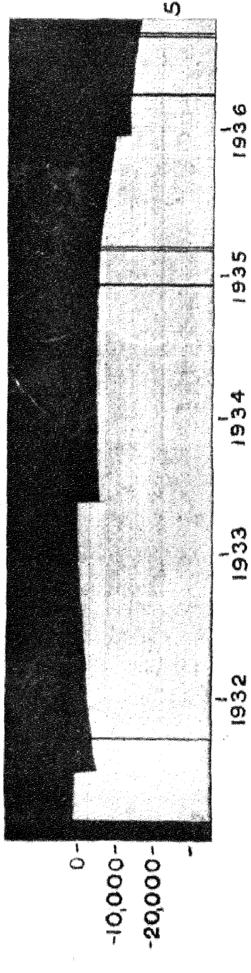


Fig. 7, June 16, 1917



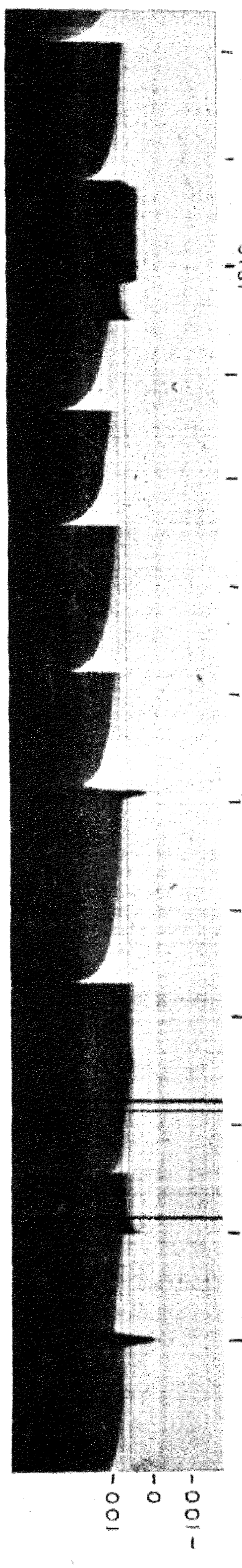


Fig. 9, May 29, 1917.

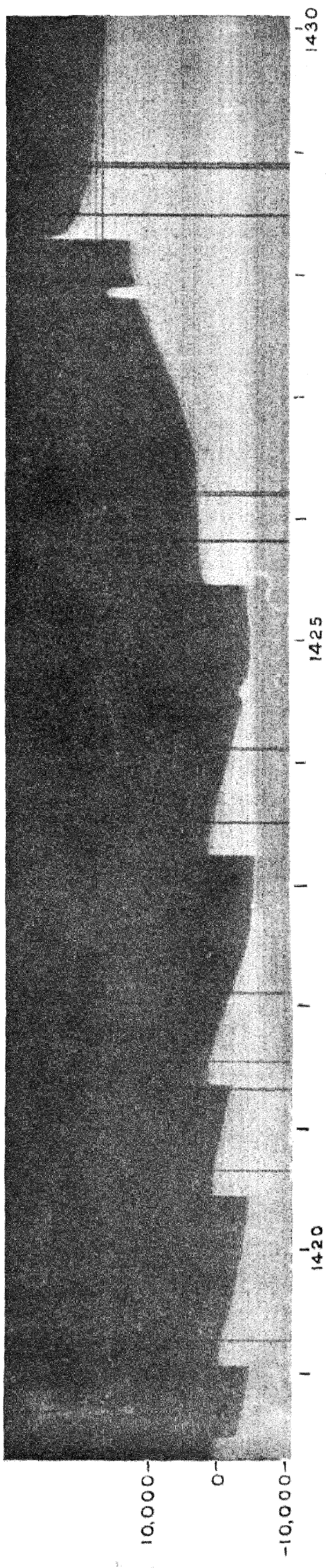


Fig. 10, August 15, 1917.

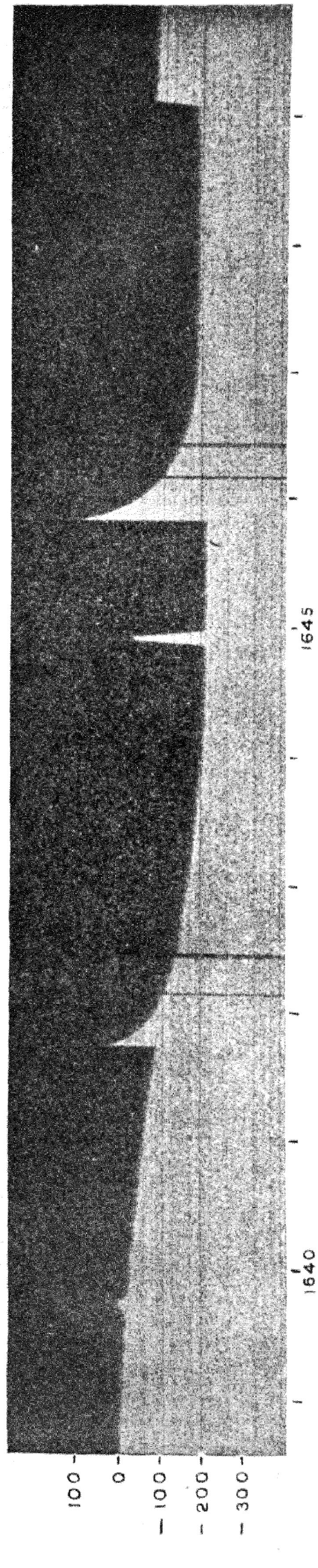


Fig. 11, June 12, 1917.

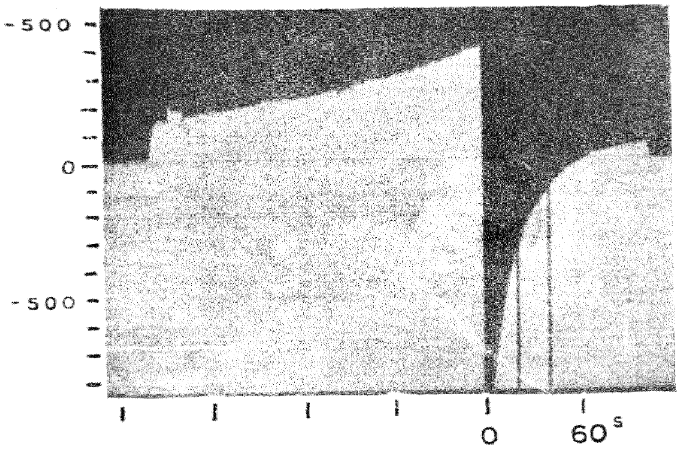


Fig. 12, May 29, 1917

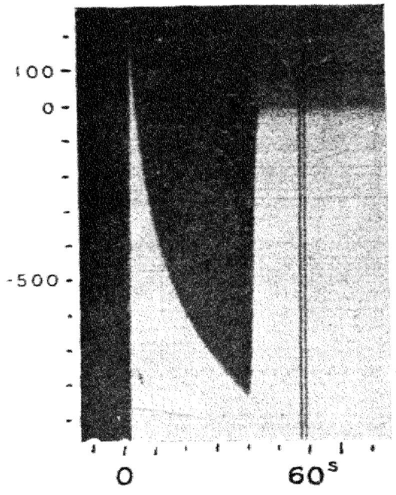


Fig 13, Aug. 15, 1917

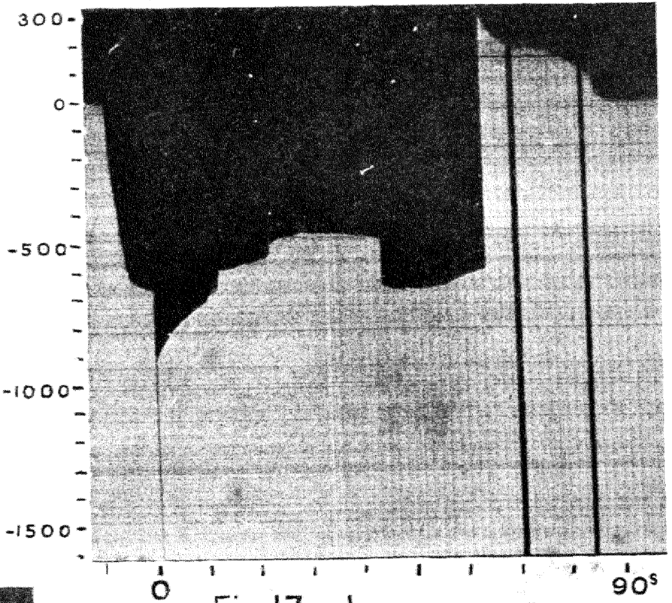


Fig. 14, June 13, 1917

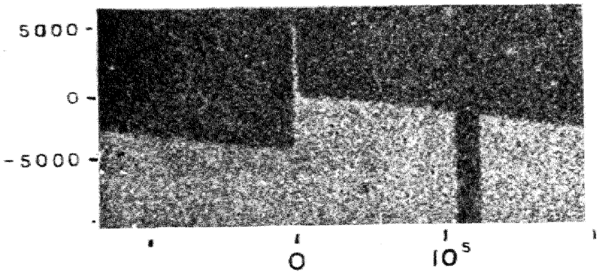


Fig. 15, August 9, 1917

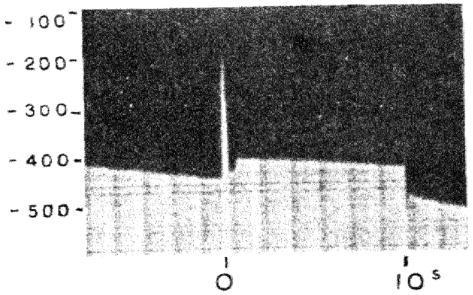


Fig. 16, June 13, 1917

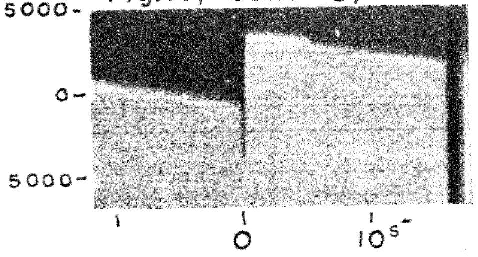


Fig. 17, June 13, 1917

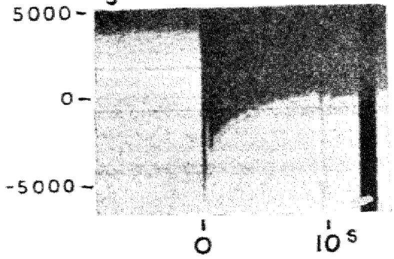


Fig. 18, June 16, 1917

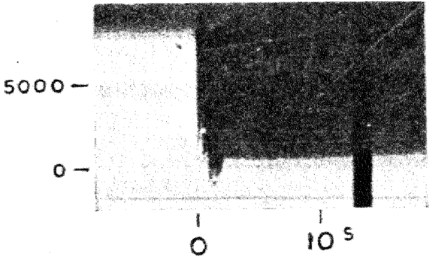


Fig. 19, June 16, 1917

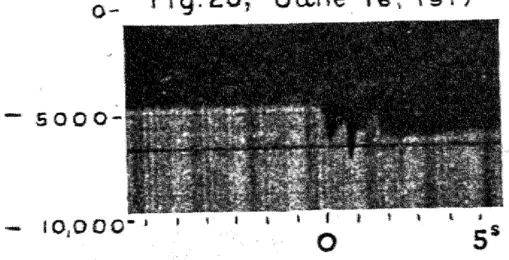


Fig. 20, June 16, 1917

of the whole tube.* If the thickness of the stratiform cloud were small, intense fields might result within the cloud and discharges might even occur; each flash would discharge only a small area of the cloud, of dimensions comparable with the thickness of the cloud.

The characteristic striated and mammatiform appearances frequently observed on the lower surfaces of stratiform clouds associated with thunderstorms may be due to intense electric fields produced as above suggested, the electrical attraction between the upper and lower charges giving rise to convection currents.

If the ionization above and below a stratiform cloud in the field of a primary thunder-cloud is unequal, the cloud will acquire unequal upper and lower charges and thus carry a resultant charge. For example, a stratiform cloud above a cumulo-nimbus cloud will intercept the flow of ions from the upper atmosphere and become charged with electricity of opposite sign to that of the upper pole of the shower-cloud, a steady condition not being reached until a potential difference between the thunder-cloud and the upper atmosphere is concentrated almost entirely in the region below the stratiform cloud. Lightning discharges between the stratiform cloud and the head of the primary thunder-cloud below will be likely to occur.

In the absence of any such cloud above the primary thunder-cloud, the great diminution of the mobility of ions or electrons dragged out of the conducting layers as they penetrate into the denser regions of the atmosphere will have a very similar effect; the concentration of charge will be greatest where the change of conductivity with the height is most rapid. We may in fact, as suggested in Section VIII., consider that the conditions are much the same as if a conducting protuberance were drawn out from the conducting layer towards the summit of the thunder-cloud. It does not seem unlikely that discharges may sometimes occur between this protuberance and the top of the thunder-cloud. In a previous paper some evidence was obtained suggesting the occurrence of discharges of very great vertical length; possibly these may have been of the type we have been considering.

---

* This action of a layer of cloud, in particular of a ground fog, in increasing the vertical electric field within it has long been recognised in the case of the potential gradient of fine weather. ELSTER and GEITEL, 'Meteor. Zeitschr.,' 17, p. 230, 1900; GEITEL, 'Physikal. Zeitschr.,' 17, p. 455, 1916.

IV. *Researches on the Elastic Properties and the Plastic Extension of Metals.*

By W. E. DALBY, *F.R.S., Professor of Engineering at the City and Guilds (Engineering) College of the Imperial College of Science and Technology.*

Original Paper received January 28,—Revised form March 24, 1920.

§ 1. *Preliminary.*

ON March 7, 1912, I described an instrument which gives photographically a load-extension diagram of a metal test piece during the process of stretching it to fracture.

On February 13, 1913, I described further experiments with the instrument.* A diagram was shown which was taken from a test piece broken in ten seconds. It is safe to say that up to that time no apparatus existed which would give a complete record of the load-extension relation during such a quick break.

I have since that time arranged the apparatus to record at even a quicker rate. Fig. 1 shows the record of a break done in 2·15 seconds. The test piece measured one inch between the shoulders. The line of the record is in dashes. These dashes fix the time scale of the diagram. Centre to centre of a pair of dashes corresponds to

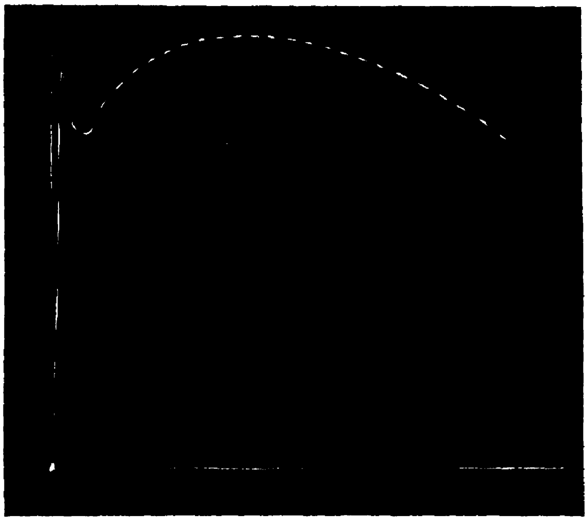


Fig. 1 (mild steel).

---

* Additional results are given in 'Transactions of the Institute of Naval Architects,' March 29, 1912, 'Institute of Metals,' May Lecture, 1917, vol XVIII., No. 2.

[Published September 29, 1920.

$\frac{1}{14}$ of a second. This time calibration is obtained by placing an interrupter in the lamp circuit. Referring to the diagram it will be seen that the yield load was reached in about $\frac{7}{14}$ seconds.

In my 1913 paper I included a diagram taken with an instrument which multiplied the extension of the gauge length 150 times so that the elastic part of the curve appeared on a scale which enabled its shape to be studied and which enabled the limit of proportionality to be identified when such a limit existed.

In my method of taking these diagrams the test piece is stretched without pause in the loading and the spot of light follows without break of continuity every phase of the relation between load and extension. Sudden slips of the crystals are duly recorded.

In the usual method the load is applied in steps, pausing at each step to observe the extension, so that the piece gets a rest under steady load during the time occupied in making the observation of extension. The load-extension curve is thus defined by a definite number of points only and peculiarities of form between these points are missed.

I have from time to time continued these elastic researches, and the following paper records some of the results obtained with what I call the *Optical Recorder of Load and Elastic Extension.*

§ 2. *The Test Piece.*

In these researches the gauge length is defined by flanges turned on the test piece itself. The ends of the arms of the extensometer rest on these flanges. The dimensions of the standard form of test piece used in these researches are shown in fig. 2. A shorter gauge length was used for the more ductile metals, but all experiments on the iron and steels were done on a 5-inch gauge length.

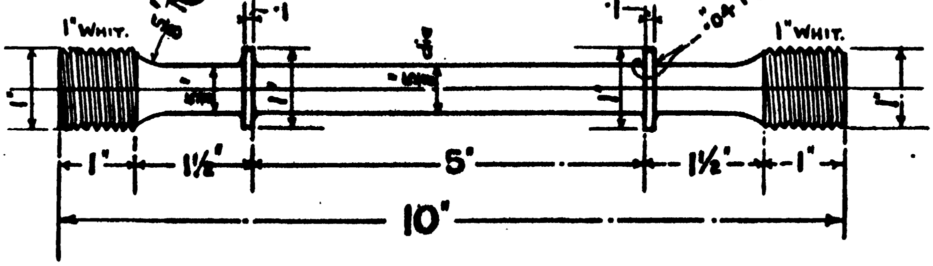


Fig. 2.

I was lead to adopt this form by the many difficulties encountered when pointed screws have to be driven into the test piece to define the gauge length.

These screws cannot be driven properly into hard material like the alloy steels which have to be tested nowadays, and in soft material like copper the primitive centre dot in which the points rest, elongates as stretching proceeds and the points slip.

The modulus of elasticity, E, found from flanged test pieces agrees with the values found from plain bars. The flanges therefore have negligible influence on the elastic extension of the gauge length. They restrict the plastic extension slightly. In mild steel the total extension is about 3 per cent. less when found from a flanged test piece than it would be if found from a plain bar.

Dr. **COKER** has kindly examined the distribution of stress produced by a flange in a xylonite test piece made to the dimensions of fig. 2. When the xylonite test piece is stretched the colours show that there is no stress in the flange itself and there is a slight but symmetrical modification of the stress distribution at its root. This means that as stretching proceeds the flange is not distorted, and therefore the distance between the flanges is a true measure of the extension of the primitive gauge length which they define.

§ 3. *The Elasticity of Materials and a Typical Load Elastic Extension Diagram of Mild Steel.*

The elasticity of a material means in a general sense its power of returning to its primitive form after loading has been applied and removed.

The recovery may be partial or complete.

The power of complete recovery is lost when the stress produced by the loading has once passed beyond a certain limiting value peculiar to the material.

Below this limiting stress the extension of a steel test piece is proportional to the load.

Above this limiting stress the extension increases at a greater rate than the load.

The limit is therefore called THE LIMIT OF PROPORTIONALITY.

The power of recovery may thus be distinguished into the power of complete recovery possessed and retained only so long as the stress in the steel has never once exceeded the limit of proportionality: and the power of partial recovery peculiar to the state into which the metal passes directly it has once been loaded beyond the limit of proportionality.

Provided that the material has never been loaded beyond its limit of proportionality the material may be said to be in a state of *perfect elasticity*, because it possesses the power of complete recovery of form after removal of load; alternatively it may be said to be in a state of *proportional elasticity* because its extension is found to be proportional to the load.

The one term includes the other. If it is found to extend proportionally to the load its recovery is perfect after removal of load.

No metal is, however, quite perfect in its recovery, but the term perfect used in the sense defined above is convenient and substantially expresses the experimental results.

A diagram recording the elastic extension of mild steel is seen in fig. 3.

This steel contains 0·156 per cent. of carbon. The extension scale of the diagram is defined by the distance between the two vertical lines seen in the diagram. This distance represents an extension of 0·01 inch.

Proportionality between load and extension ceases at about 4·5 tons corresponding to a stress of 14·67 tons per sq. inch. Yield occurs at 6·7 tons which corresponds to 21·85 tons per sq. inch. The load drops away from yield to about 5·5 tons giving a stress of 18 tons per sq. inch. These stress are reckoned on the original area of the cross-section of the test piece.

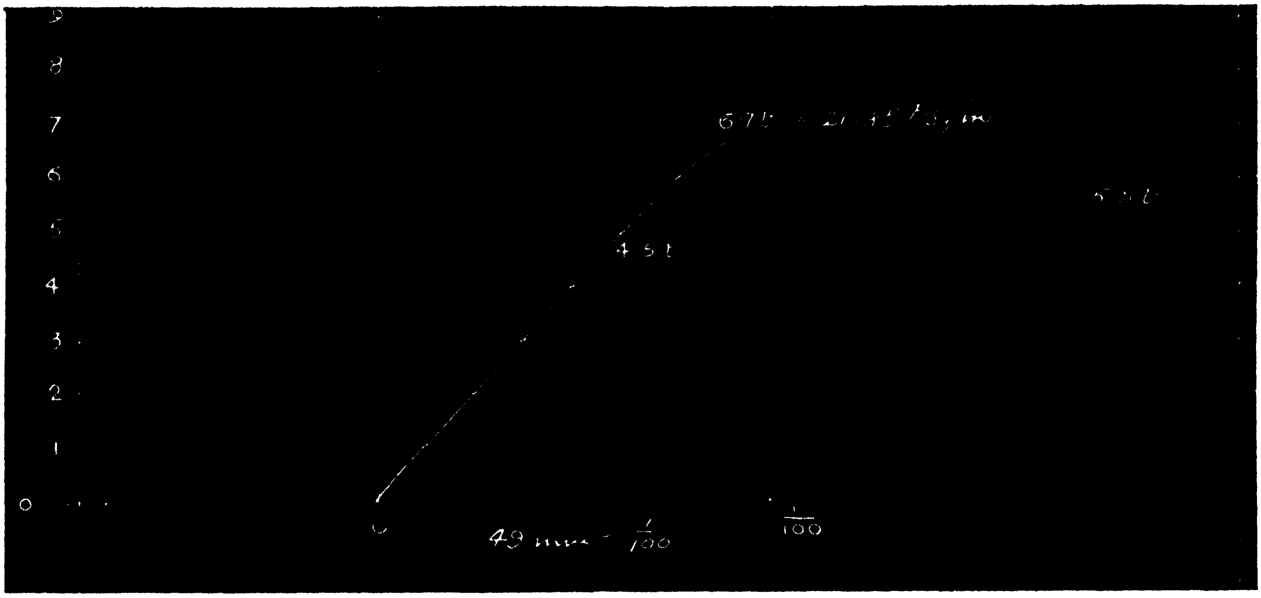


Fig. 3 (mild steel).

The slope of the line from the origin to the limit of proportionality defines E, the modulus of elasticity. From the diagram its value is 13,300 tons per sq. inch.

§ 4. *Restoration of Perfect Elasticity after Overstrain.*

The term "OVERSTRAIN" means that a metal has been loaded beyond its limit of proportionality.

If the load is removed after a test piece has been strained beyond the limit of proportionality and then the piece is immediately re-tested, the record shows a curved line.

It has no range of proportionality and no modulus of elasticity which can be identified with E.

The material still possesses elasticity because it shrinks as the load is removed, but the elasticity is imperfect in the sense that change of length is no longer proportional to change of load.

But, if the metal is iron or mild steel, proportional elasticity is slowly recovered with time; and this change from unproportional to proportional elasticity or from imperfect to perfect elasticity is accelerated by boiling.

In fact overstrained iron or mild steel is restored to its perfect or proportional elastic state with remarkable rapidity by mere boiling. This point has been established by Sir ALFRED EWING.*

I have found, however, that overstrained high carbon steels and the alloy steels do not recover proportional elasticity either by resting or by boiling.

The elastic line of a 3 per cent. nickel steel is seen in fig. 4. It is lettered A.

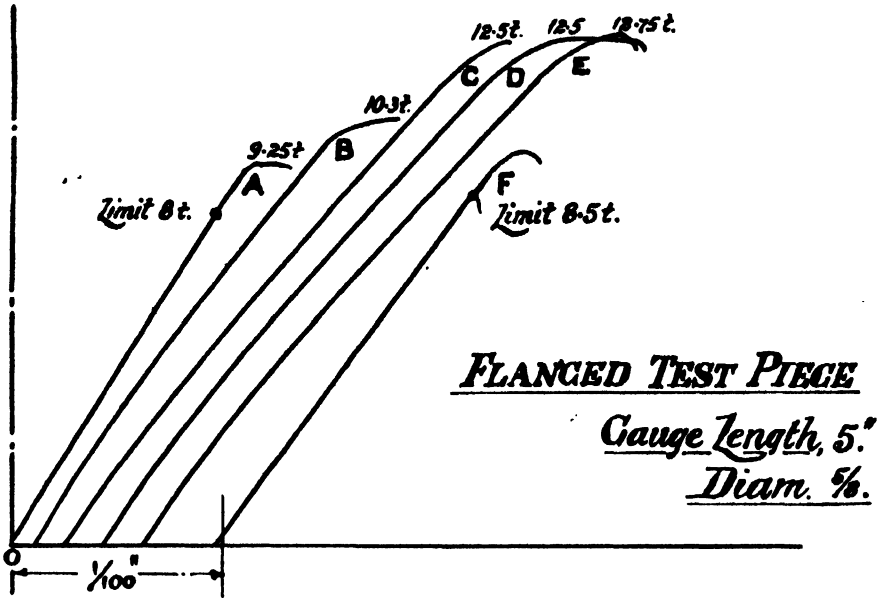


Fig. 4 (nickel steel).

The limit of proportionality is reached at 8 tons, 26 tons per sq. inch, and the yield at 9·25 tons, 30 tons per sq. inch.

The piece was stretched 2 per cent. and then the diagram, line B, was taken. Proportional Elasticity has disappeared.

Curve C is the record after a 6 per cent. stretch.

Curve D is a repetition test after turning the bar to a slightly reduced diameter. The interval of time between C and D is 24 hours. No restoration of elasticity has taken place. It has been established by other experiments that a lapse of many months has no effect in restoring the proportional elasticity.

The piece was then boiled for 1 hour, and curve E shows that elasticity has not been restored.

Finally, the piece was heated to 550° C. in a muffle furnace for about half an hour and was then allowed to cool down with the furnace. Line F, taken immediately after this treatment, shows perfect recovery of proportional elasticity and a slight raising of the limit and the yield point.

I have confirmed these results by other experiments on nickel steel test pieces and on high carbon steel test pieces.

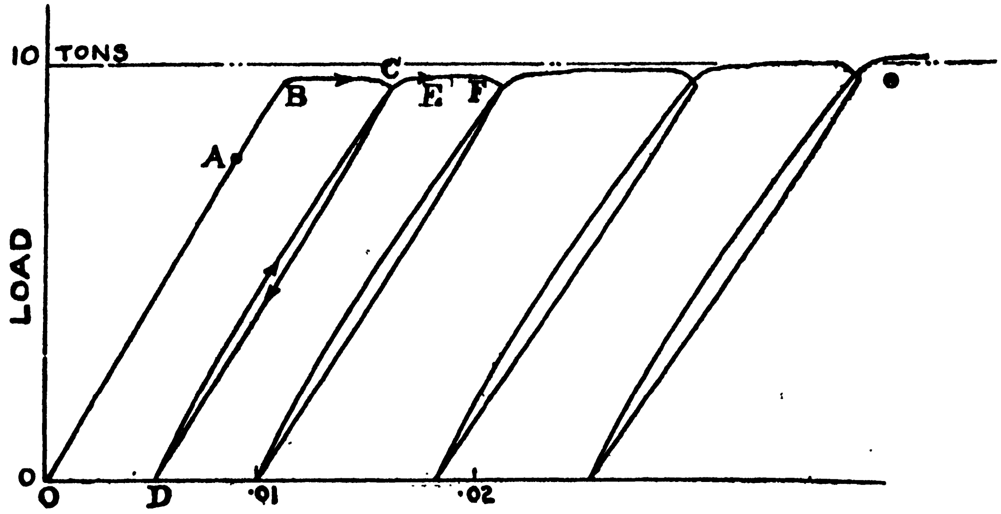
---

* 'Phil. Trans. Roy. Soc.,' 1899.

§ 5. *Looping after the Elastic Limit of Proportionality has been Passed.*

The recorder is fitted with a microscope so that the process of stretching can be watched as the experiment proceeds, and the loading, which is produced by hydraulic pressure, stopped at any moment. This arrangement enables interesting records to be taken, because after the test piece has been stretched an assigned amount, the load can be let off and then immediately re-applied, so that stretching continues through a second interval and so on.

Such a record is seen in fig. 5. The material is nickel steel. It will be seen that the removal and the re-application of the load compels the spot of light to trace a loop. The area of the loop represents the internal work done during the process.



1ST PLATE (492 A) G.L, 5", DIA. ·625".

25TH JULY 1918.

Fig. 5 (nickel steel).

Following the path of the spot it starts from the origin O and describes the elastic line OA, passes the limit of proportionality at A, and then curves away to the yield point B, and on to C. At C the loading is stopped, the hydraulic pressure is relieved by opening the exhaust valve, and the spot travels down the curved path CD as the load falls to zero. The exhaust valve is then closed and the pressure valve is opened and the process is repeated through the path EF and so on.

When the steel test piece has been stretched beyond its limit of proportionality, for example to C, fig. 5, the total extension is made up of two parts, namely :—

(1) the proportional elastic extension up to the limit of elasticity, for example up to A, fig. 5 ;

(2) the plastic extension after the limit has been passed.

When the load is removed the steel shrinks unproportionally, for example from C to D, by an amount approximately equal to its previous proportional elastic extension and then stops at a dimension greater than its primitive dimension by an amount approximately equal to its plastic extension.

The increase of size measured after the removal of the load is called the Permanent Set. OD is the permanent set produced by the first stretching of the test piece to C. From the diagram the permanent set measures 0·0052 inch and the unproportional elastic recovery measures 0·011 inch.

When the load is re-applied the spot of light moves from D to E along a curved path. The extension is no longer proportional to the load. The metal is in a different elastic state. Stretching beyond the limit of proportionality has robbed the metal of its power of proportional extension and of perfect recovery after removal of load. It may be said to be in a state of *imperfect elasticity*, or alternatively it may be described as in a state of *unproportional elasticity.*

The imperfect state is disclosed by the loop formed by the removal and re-application of the load.

The diagram shows four loops. Each loop is slightly larger than the loop preceding it.

The four loops shown were all recorded on a half plate inserted in the camera. A succession of plates was taken and the last plate is shown in fig. 6.

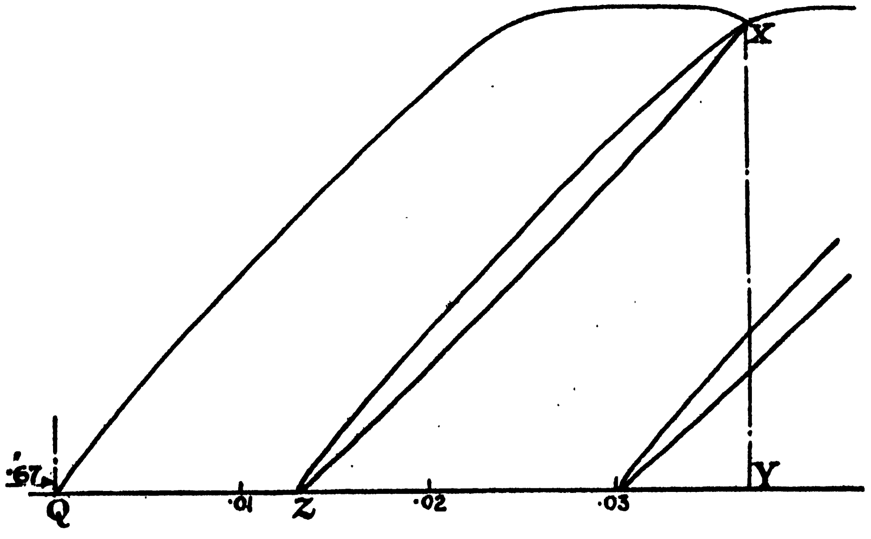


Fig. 6 (overstrained nickel steel).

This last plate shows that unproportional elastic shrinkage occurs right up to the load at which local contraction begins. The last loop is just seen on the plate.

The last line seen curving up from the origin Q is the typical curve of overstrained material. The primitive gauge length of 5 inches had been stretched to 5·67 inches before the last plate, fig. 6, was taken.

The permanent set QZ measured from the new origin Q is 0·0127 inch. The total permanent set is thus 0·67+0·0127 inch = 0·6827 inch. The unproportional elastic recovery is ZY = 0·0244 inch.

§ 6. *Looping a General Property of Metals.*

Records of looped diagrams are shown in the following figures. The load scale is varied to bring out the shape of the loops.

The extension scale is substantially, 1½ inches measured horizontally on the diagram represents $\frac{1}{100}$ inch extension of the gauge length. The diagrams are reproduced as taken, the object being to compare the elastic line and the loop formations.

Staffordshire Iron.—Fig. 7. The limit of proportionality is reached at 3·5 tons; 11·4 tons per sq. inch. The yield is reached at 5·1 tons; 16 tons per sq. inch. The load then drops to about 4·7 tons. The loop area is not large, but the area increases progressively.

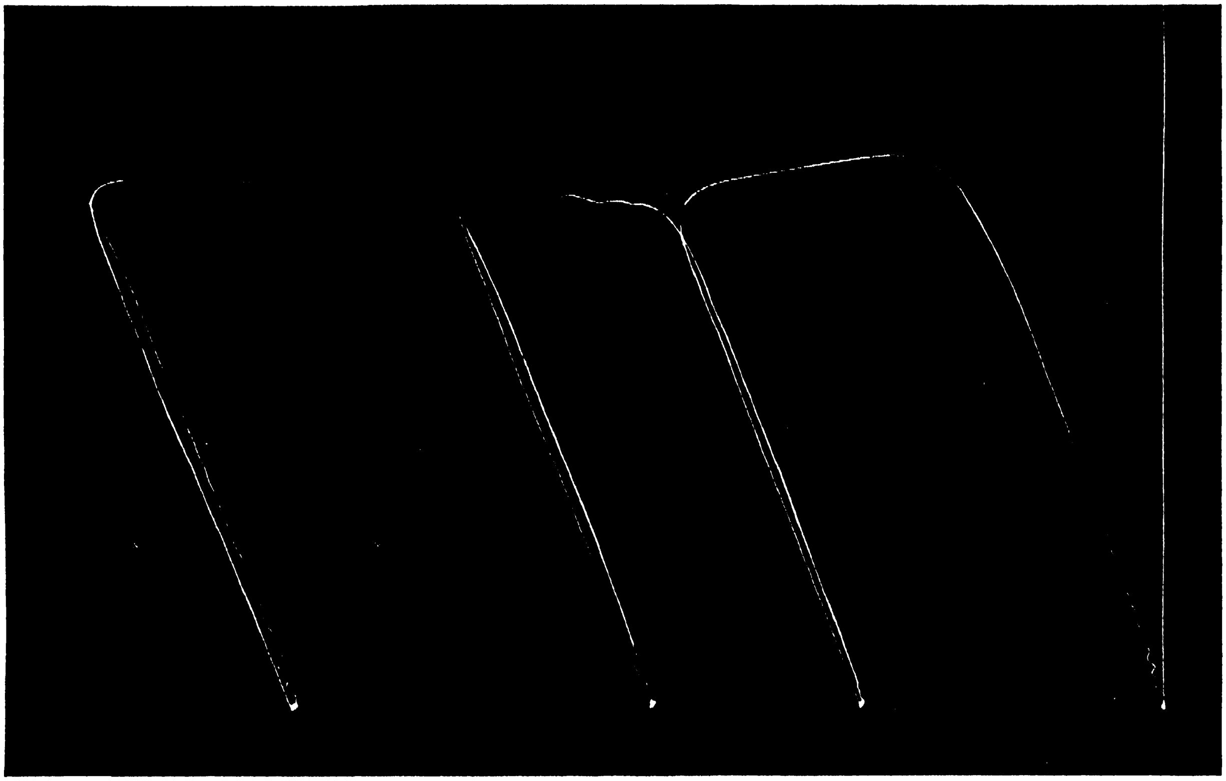


Fig. 7 (iron).

Steel.—Carbon 0·8 per cent., fig. 8. The limit of proportionality is reached at about 8 tons; the yield at 8·85 tons. There is a slight drop at the yield load. The area of the first loop is many times larger than the area of the first loop of the previous record, and this area increases rapidly in the succession of loops.

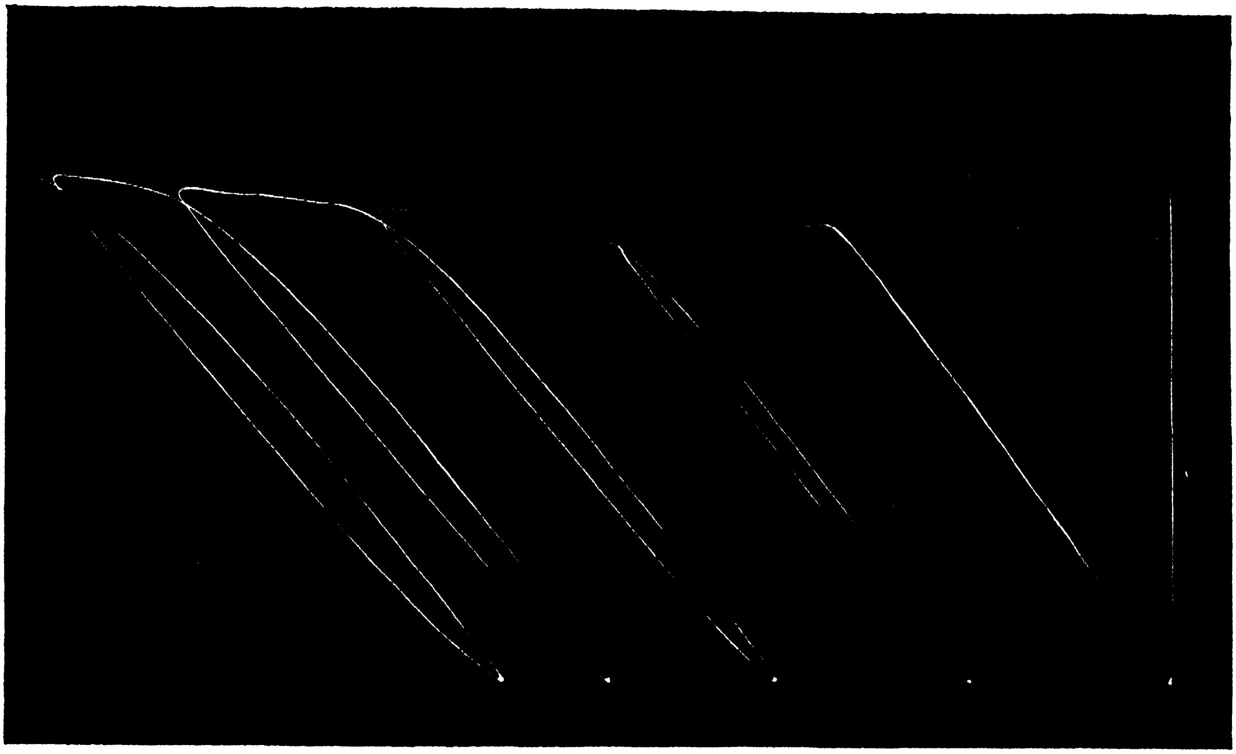
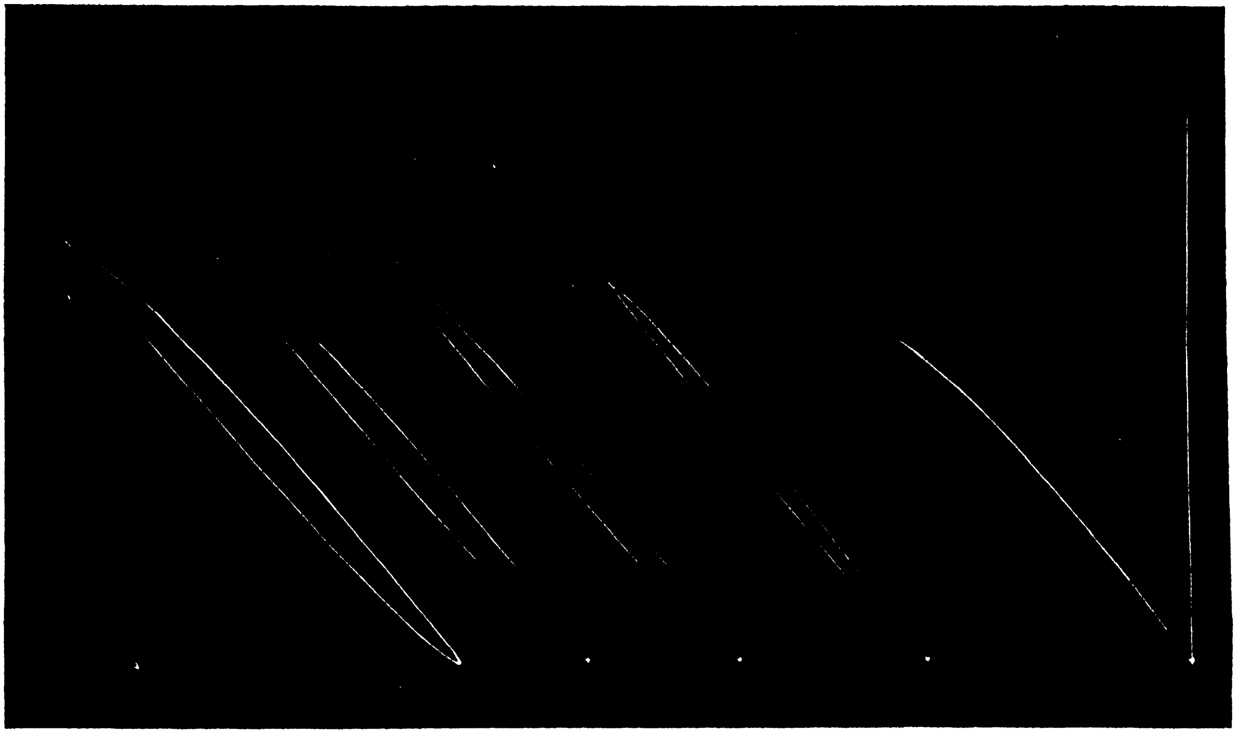


Fig. 8 (0·8 carbon steel).

Steel.—Carbon 0·8 per cent., fig. 9. This diagram is introduced because it is taken from a test piece cut from a bar delivered from the works as steel of the same kind and quality as that from which the previous diagram was taken.



The limit of proportionality occurs now at about 3 tons and there is no definite yield point. The loop areas and their rates of increase are about equal in the two plates.

The explanation of the difference in quality shown by comparing the two diagrams may be found in the fact that the test piece of fig. 8 was cut from a bar delivered before the war. The test piece of fig. 9 was cut from a bar delivered towards the end of the war. There has clearly been some change in the manufacturing process.

Nickel Chrome Steel.—Fig. 10. The ultimate strength of this steel found from a bar 1 inch diameter is 54 tons per sq. inch, with an extension of 14 per cent. on 8 inches and 55 per cent. reduction of area. The limit of proportionality is at a load of 10 tons on the standard test piece 0·625 inch diameter, corresponding to 32·5 tons per sq. inch. Yield sets in at 11 tons, that is, 36 tons per sq. inch.

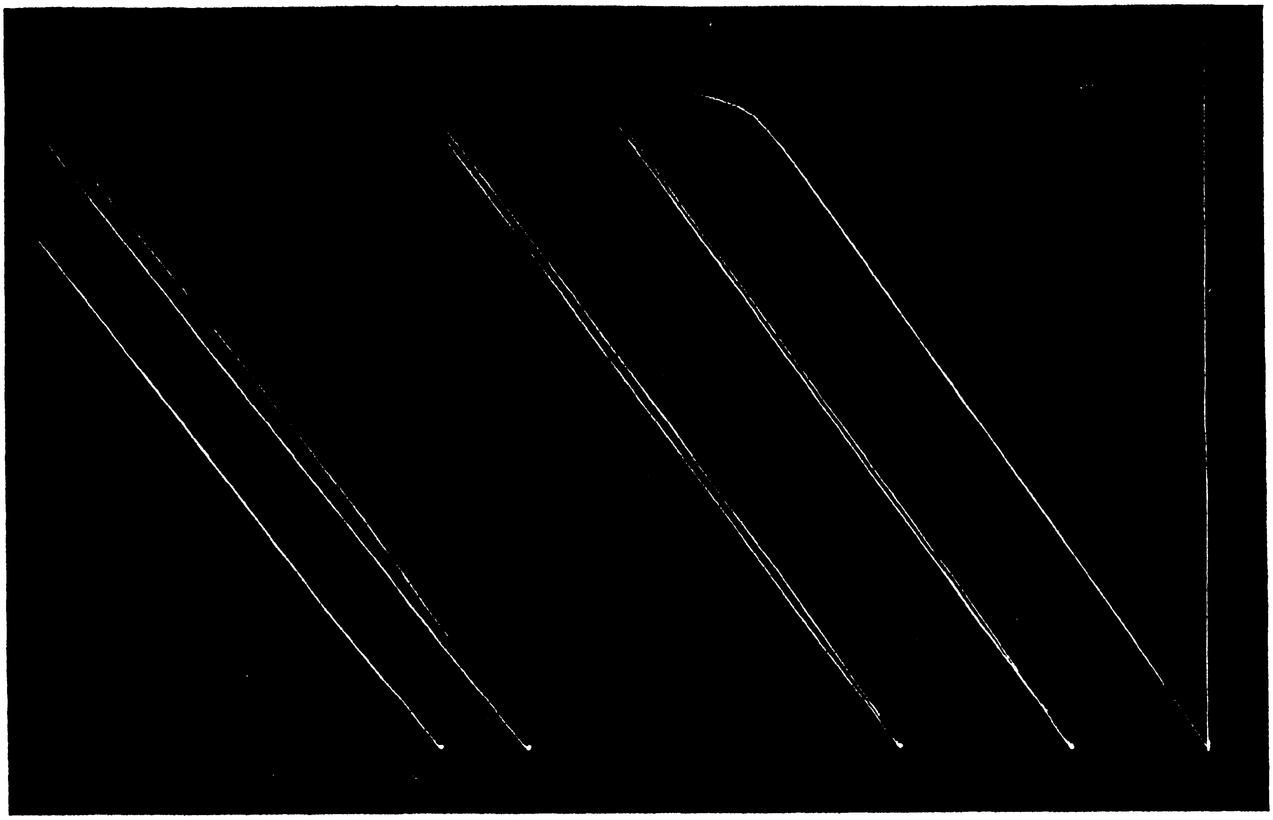


Fig. 10 (nickel chrome steel).

The first loop of the diagram is small, but the area increases rapidly, as will be seen from the three loops visible in the record.

Nickel Steel.—Carbon 0·33 per cent., Ni 3·52 per cent.—Fig. 11. The ultimate strength of this material is about 48 tons per sq. inch, with an elongation of 20 per cent. on 5 inches and a reduction of area of 47 per cent. Limit of proportionality occurs at about 30 tons per sq. inch and yield at 32 tons per sq. inch. The limit of proportionality here approaches quite near to the yield point.

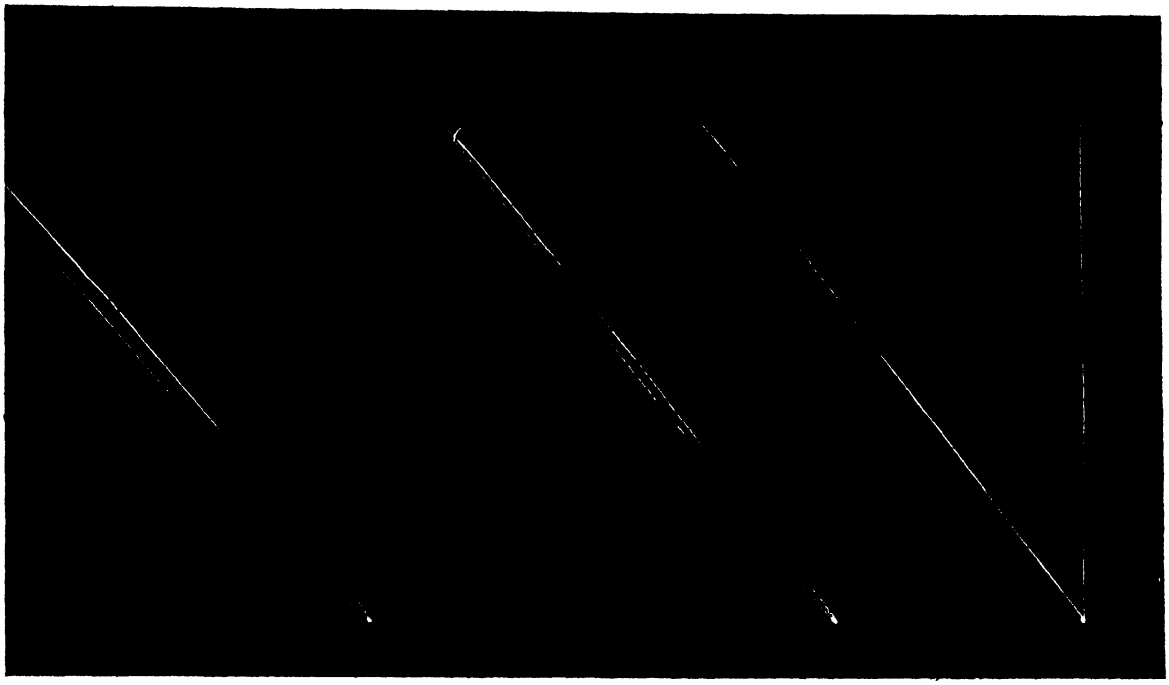
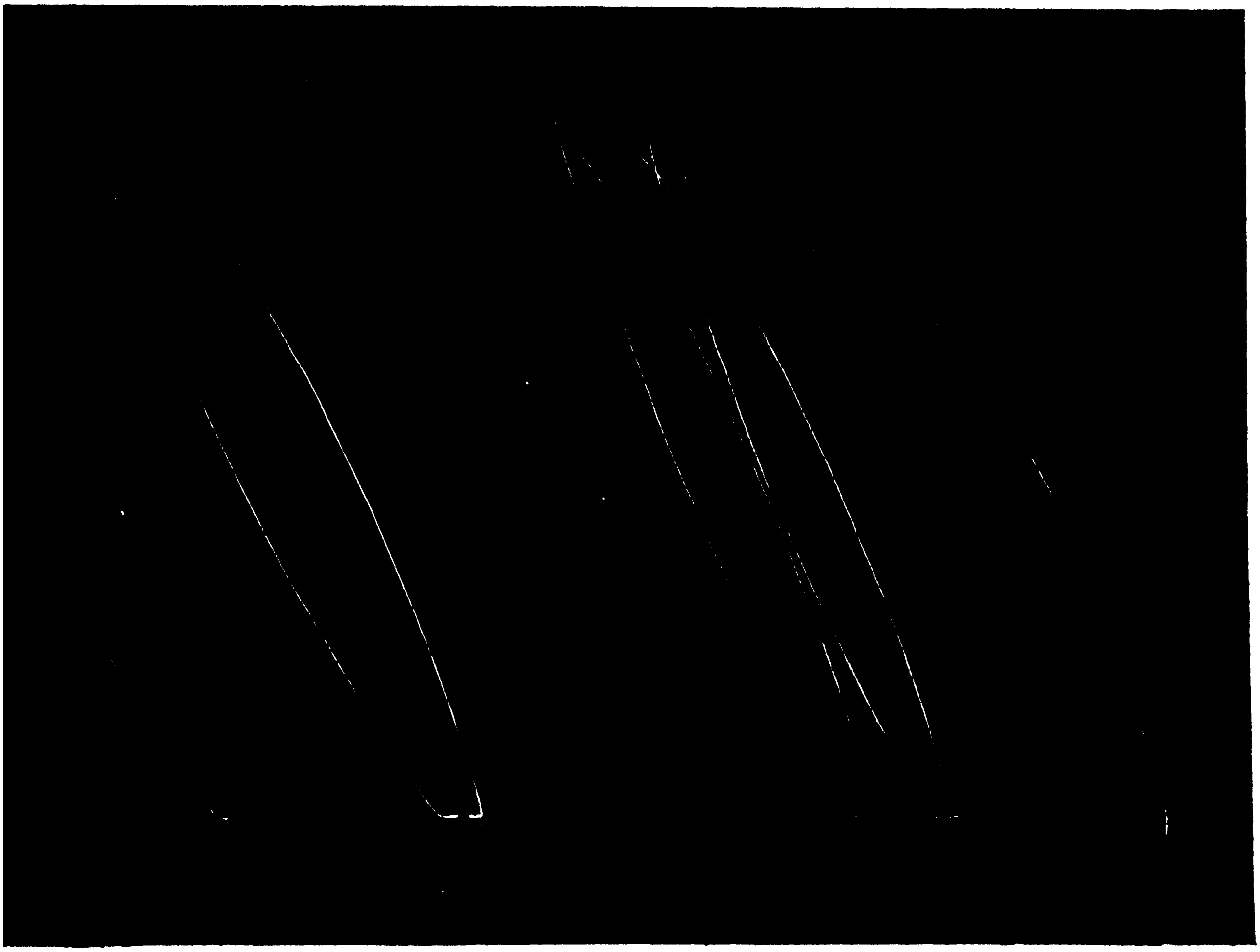


Fig. 11 (nickel steel).

Zinc.—Fig. 12. Diameter of test piece 0·8 inch. Gauge length 5 inches. The test piece was turned from a zinc rod. There is no proportional elastic line. Curvature



begins at the origin, so that the extension is increasing at a greater rate than the load from the commencement of loading. There is no definite yield point.

The most interesting result to notice is that the test piece goes on shrinking in length after the load has been removed. It is shrinking under the action of its own internal molecular forces because it is entirely free from external load.

The shrinking at no load is indicated by the flat bottom of the loop. A dwell of 1½ minutes was made in the experiment after the removal of the load and before the re-application of the load. All the perceptible shrinking at no load takes place within this time interval.

At the third loop after the load was removed the light was shut off and flashed on at intervals of $\frac{1}{10}$ seconds to get some idea of the rate of shrinking.

Tin.—Fig. 13. Diameter of test piece 0·8 inch. Gauge length 5 inches. This test piece was turned from a bar of tin. It exhibits properties similar to zinc on a smaller scale. There is shrinking continuing for about 1 minute after the load has been removed, and there is the same absence of a proportional elastic line.

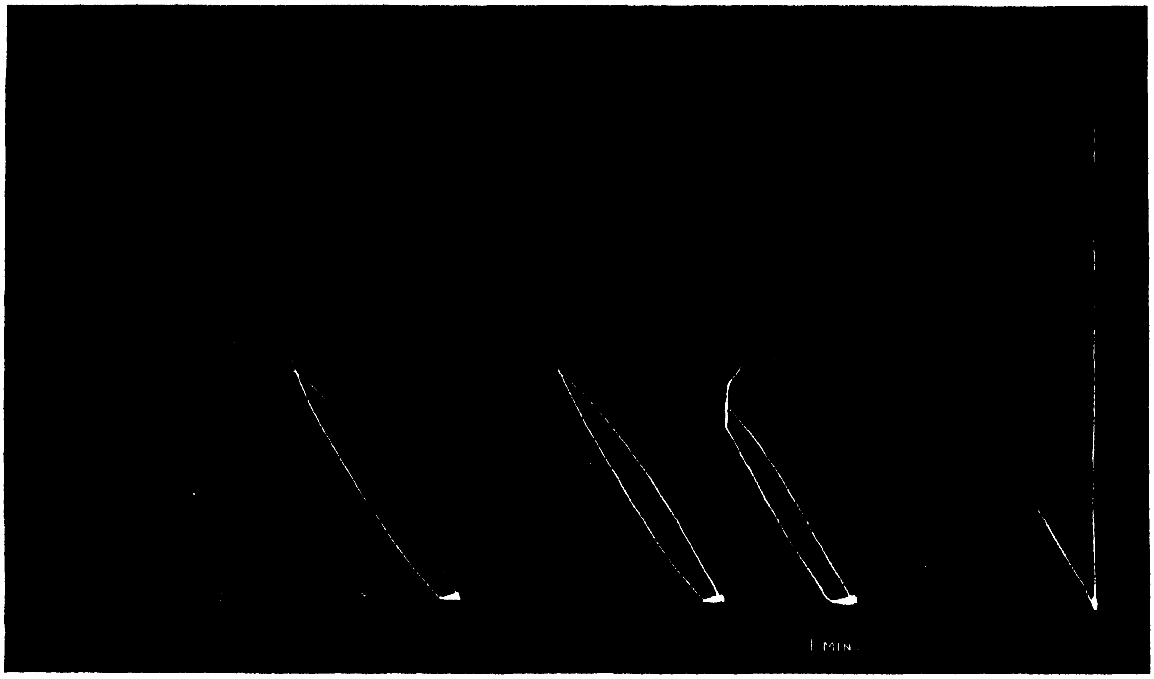


Fig. 13 (tin).

Copper.—Pure and free from arsenic. Fig. 14. Diameter of test piece 0·8 inch. Gauge length 5 inches. There is no marked limit of proportionality and no yield point in this material. The noteworthy feature of the record is the small rate of increase of loop area. This small rate of increase of loop area is a common characteristic of all the copper samples which I have tested. It may be that this rate of increase is identified with the quality of toughness.

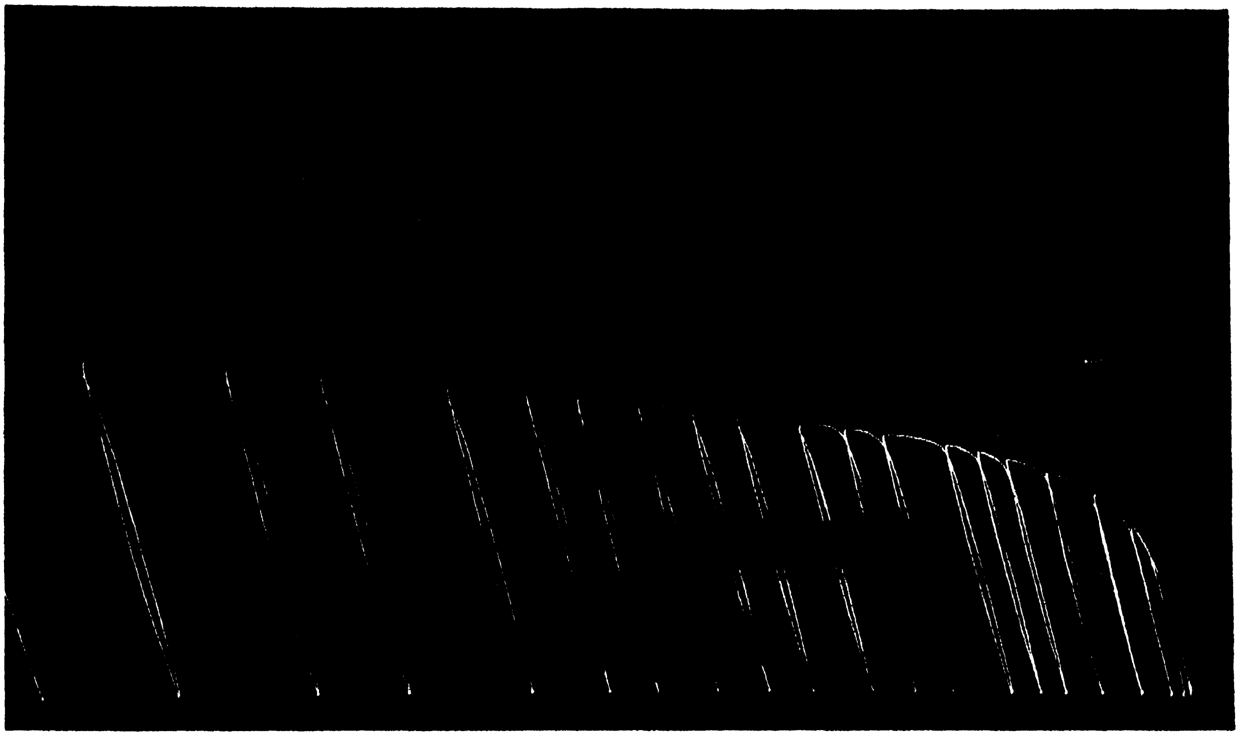


Fig. 14 (electrolytic copper).

Copper.—Fig. 15. 99·4 per cent. copper. Arsenic present, and by difference estimated at 0·4 per cent. Diameter of test piece 0·8 inch. Gauge length 5 inches.

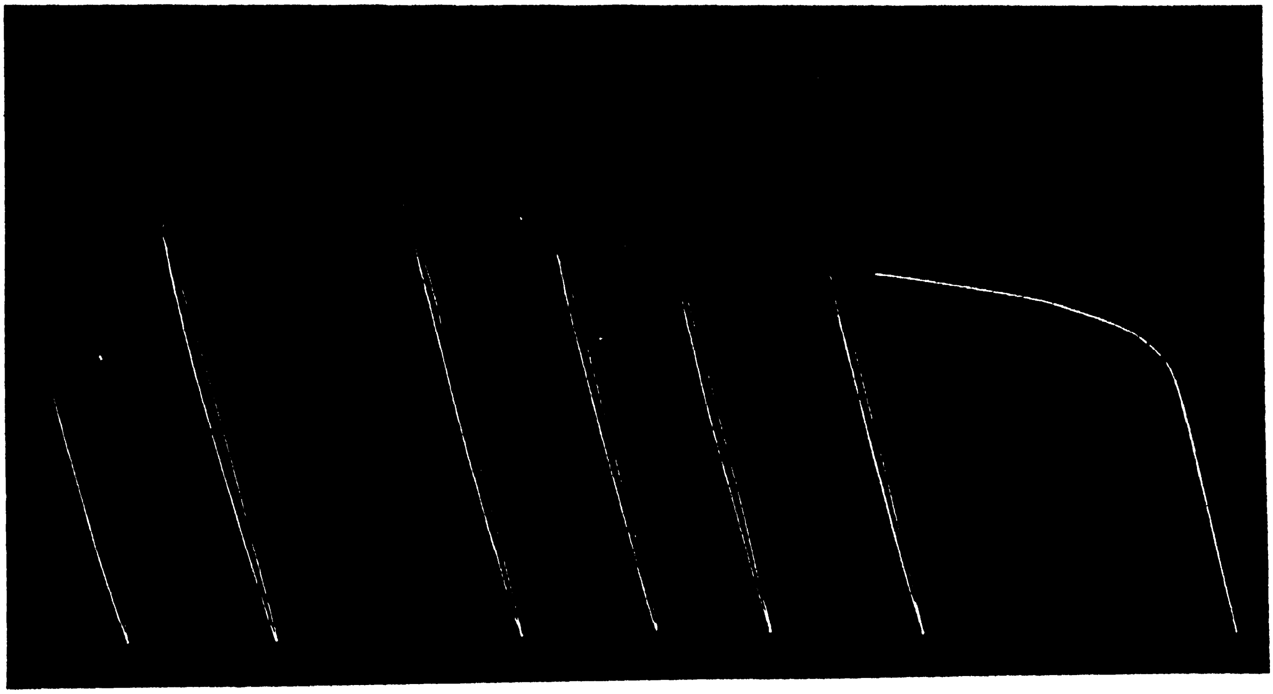


Fig. 15 (arsenical copper).

The effect of the arsenic is remarkable. It gives to the copper an elastic line with a distinguishable limit of proportionality of 1·4 tons; 2·8 tons per sq. inch.

There is no definite yield point. The loops are small and the rate of increase of loop area is small.

The elastic line from the origin to 1·1 tons is thicker than the continuation of the line. This thickening is brought about by a removal and a re-application of the load. The spot of light travelled three times up and down this piece of the diagram, indicating that the elastic line, within the limits of this load, is permanent.

Brass.—Fig. 16. Composition 60 per cent. copper, 40 per cent. zinc, with traces of tin and other impurities. The ultimate strength of the material is 32·6 tons per sq. inch. There is a marked limit of proportionality at $2\frac{1}{4}$ tons; 7·33 tons per sq. inch. Yield follows gradually. There is no contraction after the load has been removed, although the material contains so much zinc.

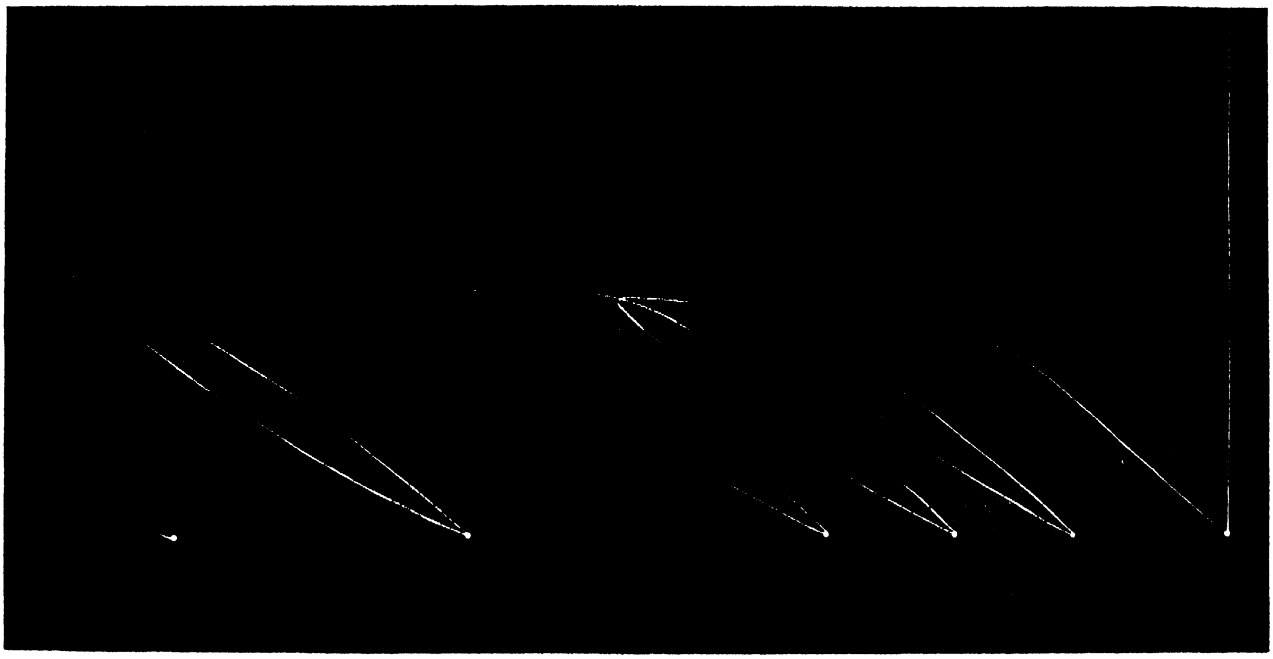


Fig. 16 (brass).

Phosphor Bronze.—Fig. 17. The curve in this diagram shows a limit of proportionality at about 2 tons; 6·5 tons per sq. inch; but it is difficult to locate the exact spot at which the line begins to curve away from the primitive straight element.

Aluminium Alloy.—Fig. 18. Diameter of test piece 0·625 inch. Gauge length 5 inches. This diagram is remarkable in that the removal and the re-application of the load in the plastic state shows no looping and therefore no hysteresis loss which can be calculated from the loop area. It appears as though the metal continually anneals itself at ordinary temperatures as plastic stretching proceeds. The material appears to be elastic up to a load of 4½ tons. The thick line indicates the remova and the re-application of load before the metal begins to yield plastically.

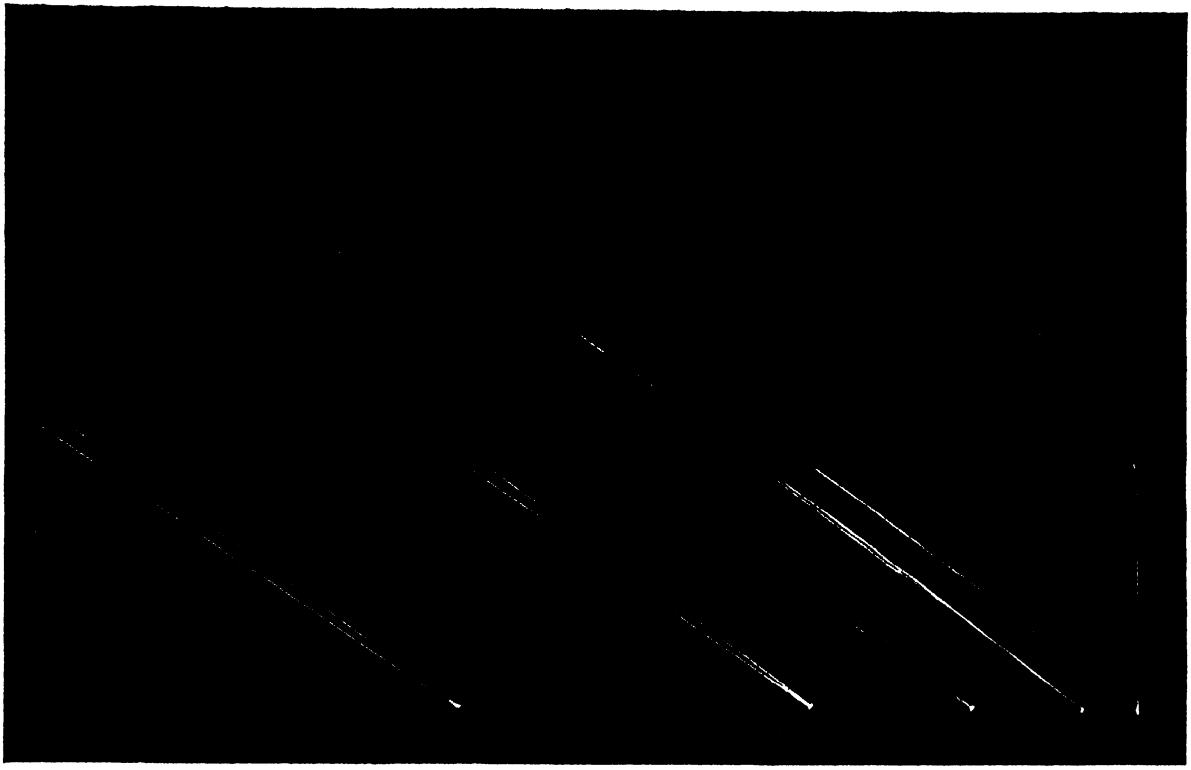


Fig. 17 (gun metal).

Fig. 18 (aluminium alloy).

§ 7. *Loop Area and Permanent Set.*

The loop area increases in size as the stretching proceeds and the rate of increase differs in different materials.

The question now arises: does the increase in area follow a regular law? The answer is given by the curves on Sheet 1 (folding diagram).

The co-ordinates on Sheet 1 are loop area and permanent set. Curve 1 shows the results obtained from a test piece of 0·8 inch carbon steel, ⅝ inch diameter, with a 5 inch gauge length. The slope of this curve shows the rate of increase of loop area as stretching is continued. The curve ends when local contraction begins. Similar curves are given on Sheet 1 for nickel steel, mild steel, and for iron.

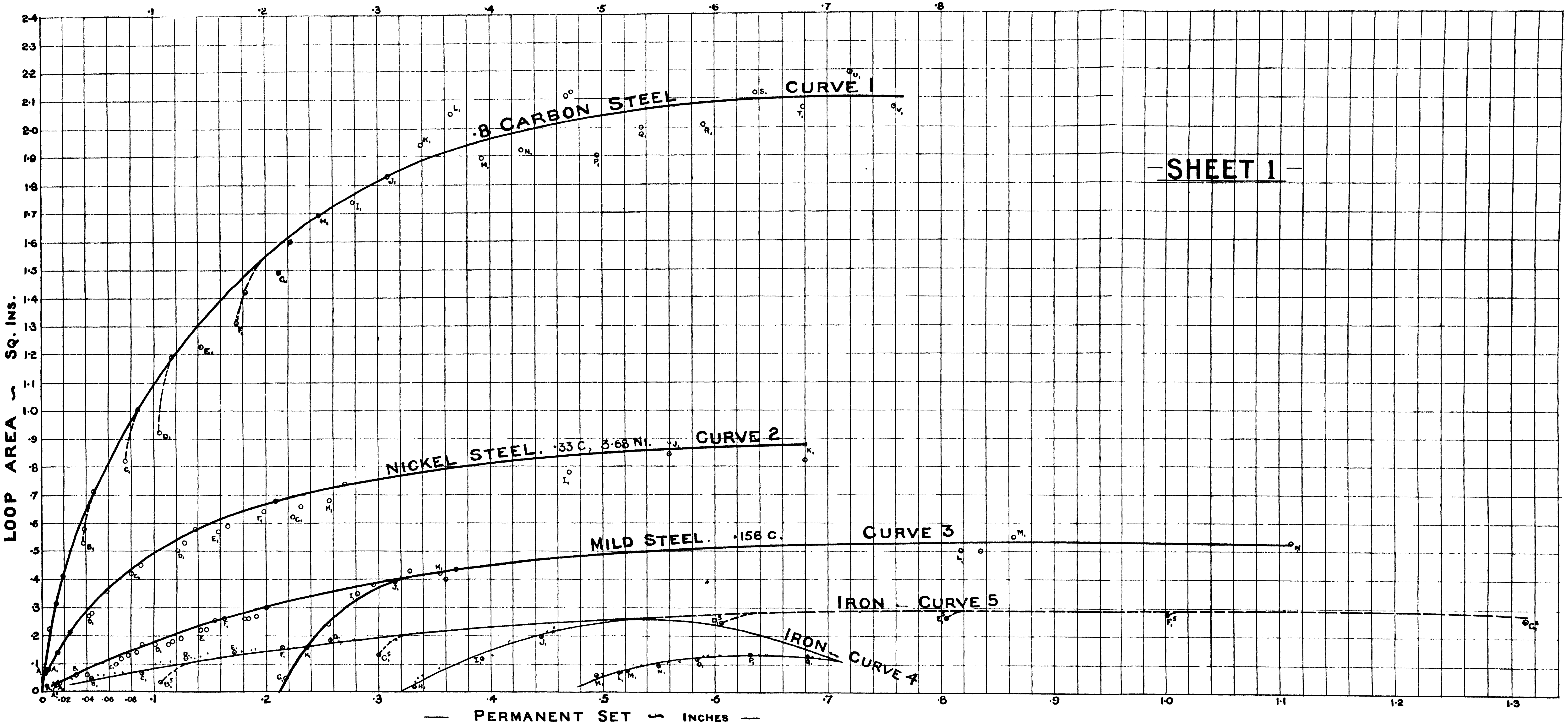
The rate of increase depends upon the time interval between the drawing of the loops and upon the kind of material. In irons and mild steels the influence of time is profound. In the alloy steels tested and in high carbon steels the influence of time is small.

When the stretching of iron or mild steel is resumed after a rest, the loop area, at first small, increases rapidly towards the area the loops would have had if stretching and looping had been continued without resting. Anticipating the detailed description of the curves on Sheet 1, this point may be illustrated by curve 3, same sheet. Plates A_1, B_1, C_1, D_1, E_1, and F_1 were taken consecutively, there being no more time interval between the plates than the few seconds required to change the plates. After the mild steel test piece had been stretched to 0·2 inch it was taken out of the machine and laid aside for 15 days. Plate G_1, the first plate taken after the rest, and plates H_1, I_1, J_1 furnish loops of rapidly increasing area until the area is reached on plate K_1, corresponding to continuous stretching without rest intervals.

§ 8. *Loop Area and Permanent Set Curve. Curve* 1 *Sheet* 1 (0·8 *per cent. Carbon Steel*).

The detailed consideration of this curve will show how all the curves of the diagram on Sheet 1 have been derived. The capital letters along the top of the procession of loops seen in fig. 19 refer to the sequence of negatives recording the loops taken from a standard test piece of 0·8 inch carbon steel.

Plate A gives the record of the first application of the load to the test piece and its immediate removal and re-application four times. The plate therefore shows the elastic line and the first four loops. A *scale* is placed under the loops so that the *permanent set* of the primitive 5-inch gauge length can be read at any point in the procession of loops. For example, the permanent set at the end of the looping operations recorded on the sequence of plates A, B, C, D, is the distance Ok = 0·137 inch.



— LOOP AREA PLOTTED AGAINST PERMANENT SET

Each small circle denotes a loop, and the letter written against some of them identifies the first loop on the plate corresponding with the letter. The subscript 1 means the 1st loop.

The procession is formed by setting the plates in due sequence and placing the origin of the record on each plate at the point on the extension scale corresponding to the permanent set measured from the test piece itself. For example, direct measurement of the gauge length after taking plate D shows that the permanent set is 0·137 inch. The origin of the record on plate E is then located at 0·137 inch on the scale.

The procession of loops seen in fig. 19 is reproduced on a small scale in order to present to the eye the complete record on a reasonably sized sheet.

Selected loops from the procession are shown full size in figs. 20 to 23. The elastic line and the first and second loops are seen in fig. 20. The area of the first loop represents an energy loss of 0·42 ft. lbs., and of the second loop 1·15 ft. lbs. The corresponding permanent sets are 0·002 inch and 0·007 inch.

The areas through the sequence of plates A, B, C, D, E, F increase gradually. The last loop of this sequence is seen in fig. 21 and it represents an energy loss of 7·42 ft. lbs. The time occupied in taking these six plates was 23 minutes.

After taking plate F the experiment was stopped. The test piece was removed from the machine and laid aside. After six days' rest it was put back into the machine and looping continued.

The first line after the rest and the first loop are seen in fig. 22. The first line shows no elastic recovery and the six days' rest has had no perceptible influence on the loop area, which represents 7·78 ft. lbs. The area is what it would have been if there had been no interruption of the experiment for a period of rest. The sequence of plates G to V was taken in 1 hour 20 minutes. The record on the last plate is seen full size in fig. 23. Stretching was stopped because local contraction had set in.

The areas of the loops in the sequence are plotted against permanent set in curve 1, Sheet 1. Each small circle denotes a loop, and the letter written against some of them identifies the first loop on the plate corresponding with the letter.

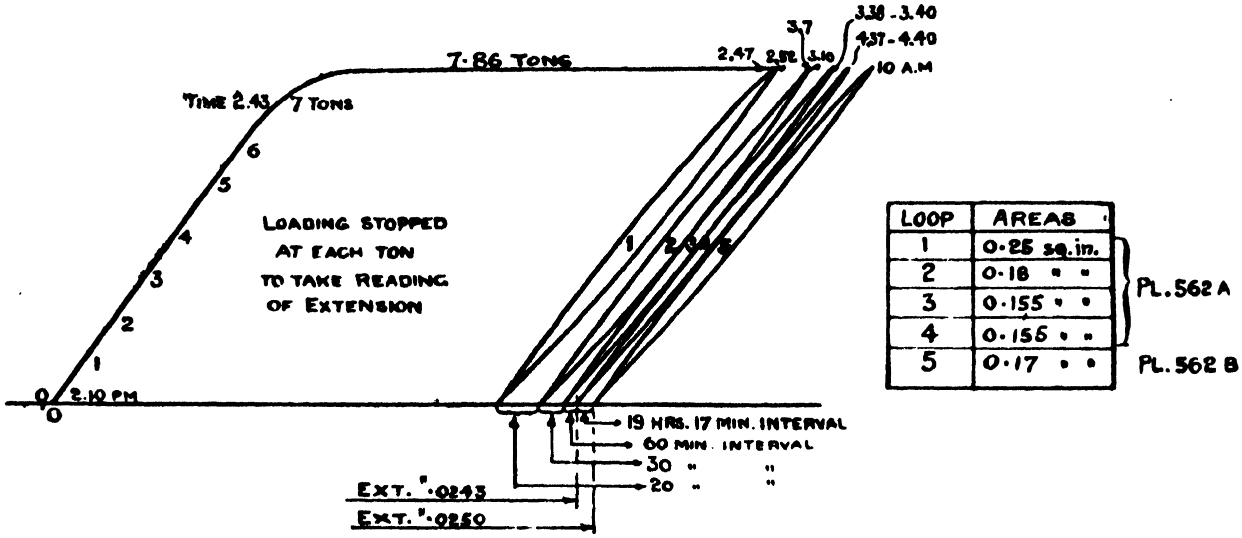
It will be noticed on Sheet 1 that the curves joining the loop areas on any one plate do not merge into one another to form a continuous curve. The time interval required to change the plate and to resume loading seems to be occupied by the material in some inner process which tends to slightly reduce the area of the next loop taken. But whatever the inner process may be, it practically exhausts itself in a few moments and produces only slight effect on the next loop area. No further change takes place after a rest of six days, and the inner process, whatever it is, has no influence in restoring the material to a state of perfect elasticity after the overstrain.

A curve sketched through the group of loops on each plate is continuous and clearly shows that the area of the loops tends to a maximum. The maximum value in this experiment represents an energy loss of 11·48 ft. lbs. per loop. This loss corresponds to 7·51 ft. lbs. per loop per cubic inch of material in the primitive gauge length.

Processions of loops were similarly taken from a nickel steel test piece, from a mild steel test piece, and from an iron test piece.

§ 9. *Looping under Constant Load.*

The diagram in fig. 24 shows the effect of looping under constant load.



LOOP	AREAS
1	0·25 sq. in.
2	0·18 " "
3	0·155 " "
4	0·156 " "
5	0·17 " "

— LOOPING UNDER CONSTANT LOAD FIG. 24

— 0·8 CARBON STEEL —

The test piece was placed in a Buckton testing machine and loaded gradually until yield began at 7·86 tons, and then loading was stopped. The extension was allowed to proceed under this load for 4 minutes. Then the load was removed and re-applied and the spot of light traced out loop 1.

Extension continued slowly under the load, still maintained at 7·86 tons, and loops 2, 3 and 4 were taken at intervals of 20, 30 and 60 minutes respectively.

The next interval was 17 hours 20 minutes, during which time the gauge length extended $\frac{1}{1000}$ inch approximately. Loop 5 was then taken. Comparing these loops it will be seen that there is no recovery of proportional elasticity although the piece was allowed to stretch under the yield load of 7·86 tons for about 20 hours.

This shows that if the yield load is kept on until the test piece has stopped extending, a process which may take a long time, at the end of the experiment the test piece will not have gained proportional elasticity. It is still in a state of imperfect, or unproportional elasticity.

§ 10. *The Practical Utility of the Load Elastic Extension Looped Diagram.*

A diagram showing the elastic line and a few loops is of great practical utility in industrial applications. The data immediately measurable from the diagram are :—

(1) the load at the limit of proportionality ;

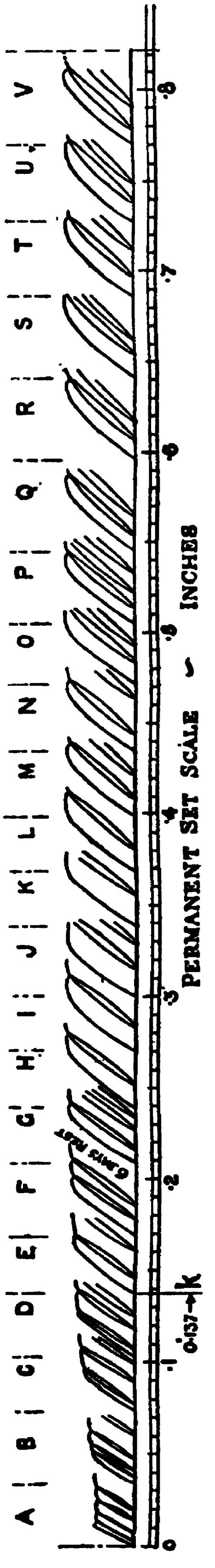


FIG. 19

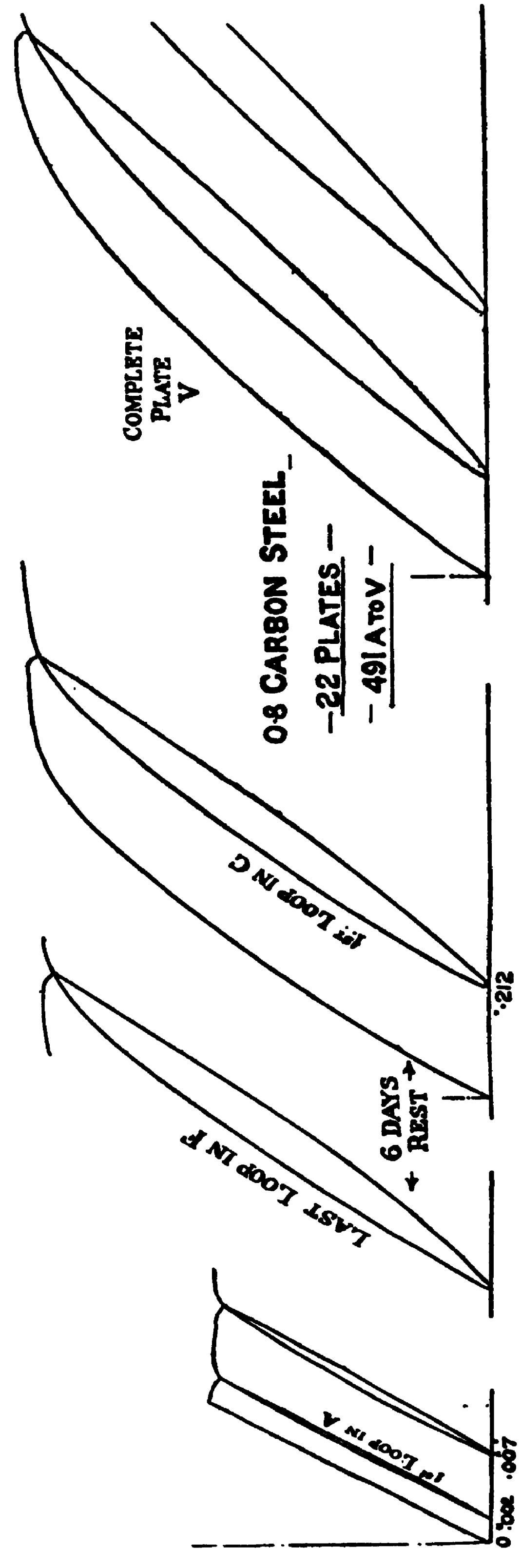


FIG. 20

FIG. 21

FIG. 22

FIG. 23.

0·8 CARBON STEEL

— 22 PLATES —

— 491 A TO V —

(2) the yield load;
(3) the value of E: this is given by the slope of the elastic line;
(4) the work lost per loop;
(5) the rate of increase of the work lost per loop.

The diagram from *normal material* satisfying known conditions of composition and manufacture may be used as a diagram of comparison. The form of the curve, the loop area and the rate of increase are sensitive to changes in the kind of material and to changes in the inner state of materials.

The diagram is specially useful in showing the load at the limit of proportionality, for this load bears neither a constant relation to the yield point, when there is one, nor to the ultimate load. Consequently factors of safety reckoned against either the yield load or the ultimate load are ambiguous.

This is specially important in gun design. The whole theory rests upon the elastic property of the material, and the theory ceases to apply after the limit of proportionality is passed.

Considerable research in many directions is necessary before a full interpretation can be given to the looped diagrams, and for the present I will reserve further discussion.

§ 11. *Correlation of Diverse Tests by the Load Extension Diagram.*

Load extension diagrams of the kind shown in this and former papers are likely to be useful to the engineer and metallurgist in the correlation of the many different tests now made to ascertain the quality of metals.

For example, fig. 25 was taken from a test piece of material giving a low impact number. Its shape differs markedly from the shape of a normal diagram. It corresponds in fact with the shape of the curve found from overstrained material.

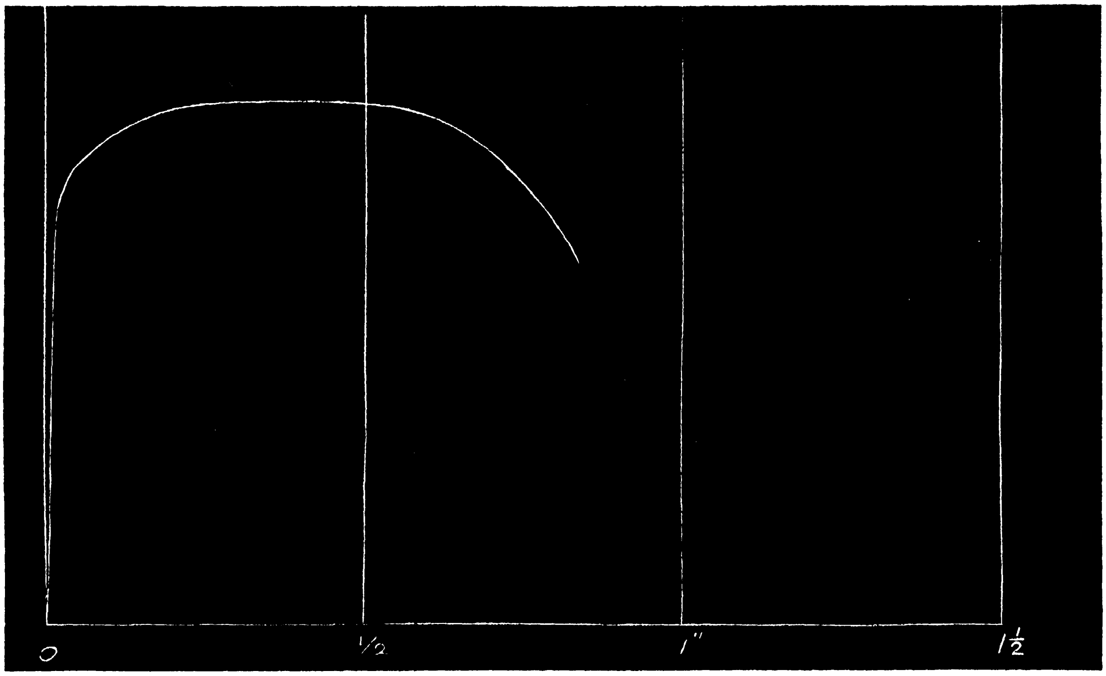


Fig. 25 (overstrained mild steel).

The inference is that the material is in the overstrained condition. That this inference is correct is shown by fig. 26. A test bar of the same material was annealed by heating to 550° C., and then cooling slowly within the furnace, and then it was found that it gave the diagram of fig. 26, which is the normal shape for the class of steel tested.

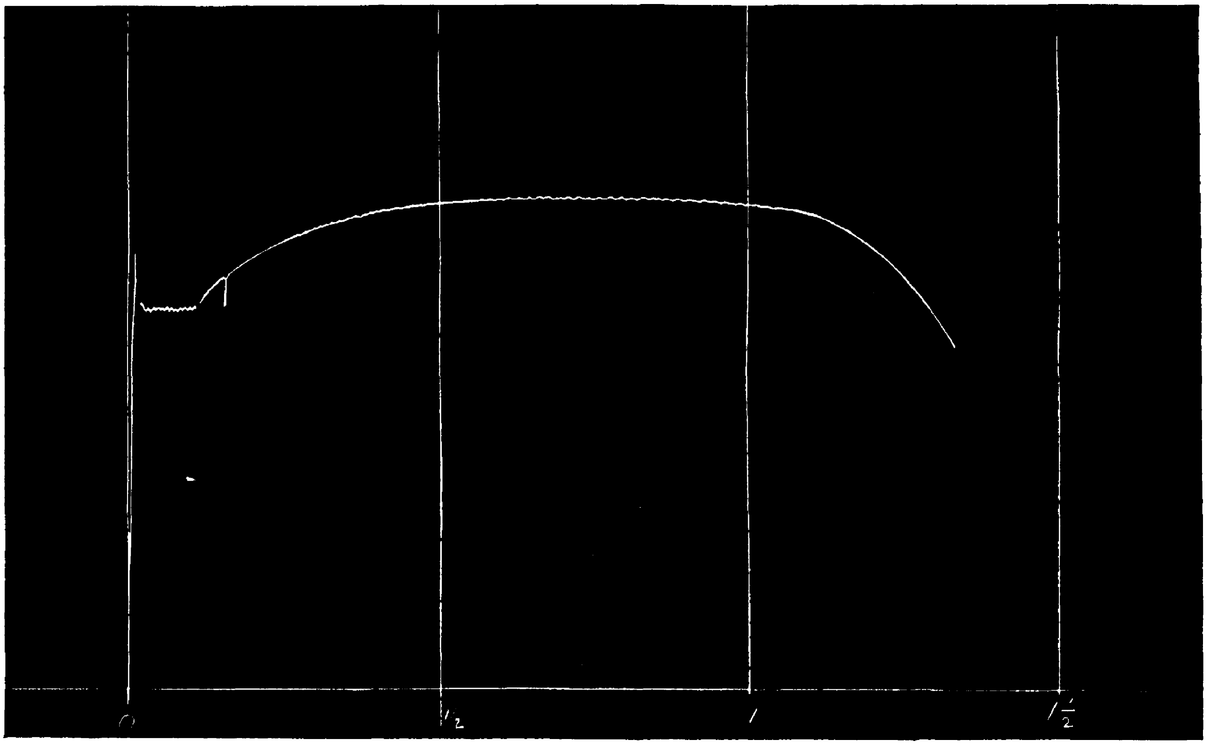


Fig. 26 (normal mild steel).

I have tested many bars of steel rejected on shock test and giving low impact numbers, and I have always found that the shape of the load extension diagram discloses the abnormal state of the metal. Much work of a comparative kind must be done before the result can be widely generalized. The War Committee of the Royal Society did a considerable amount of work in this direction.

The limiting fatigue stress may probably be found by inspection from a load elastic extension diagram.

It is probable that *the limiting range of stress in fatigue has for its positive value the stress equal to the limit of proportionality.*

Referring to the diagram for iron, fig. 7, it will be seen that the limit of proportionality is at about 3½ tons, corresponding to 11·45 tons per sq. inch.

I prepared six test pieces of the material, and Dr. Stanton kindly applied his fatigue test to them at the National Physical Laboratory. He found, after applying alternating loads in the aggregate 24,000,000 times to the eight test pieces, that the approximate limiting range of stress in fatigue was between ±10½ and ±13 tons per sq. inch. The average is ±11·75 tons per sq. inch.

The agreement between the limit of proportionality shown on the diagram, namely, 11·45 tons per sq. inch, and the fatigue limit found by quite a different test, is remarkably close.

Again, if this result could be generalized, it could be asserted that every load elastic extension diagram shows the positive value of the fatigue limit. The long and tedious experiments with alternating loads would be unnecessary. Such a conclusion requires comparison to be made over a wide range of material. This, therefore, is a promising field of research.

These brief notes of results and inferences show what a wide range of information lies before the engineer and metallurgist if he has before him a pair of diagrams, the one showing the complete load extension curve from zero to fracture, the other a load elastic extension looped diagram on a large extension scale.

The matter incorporated in this paper has been selected from experiments extending over several years. I desire to express my acknowledgments and thanks to Prof. WITCHELL for his help, and in particular for the assistance he gave me in reducing the looped diagram to the curves of Sheet 1.

I also desire to acknowledge the assistance of Mr. ORR for the skill and care with which he has drawn the curves on Sheet 1 and figures 19–23 from the photographic plates.

V. *The Stress-strain Properties of Nitro-cellulose and the Law of its Optical Behaviour.*

By Prof. E. G. COKER, *F.R.S., and* K. C. CHAKKO, *M.Sc., University College, London.*

Received September 22, 1919,—Read January 22, 1920.

THE physical characteristics of transparent bodies capable of resisting stress have been the subject of much investigation, and in particular the properties of various glasses* have been studied with much thoroughness since these latter have an extensive use both for commercial and scientific purposes.

In recent years many new forms of optical materials have found an industrial use, and especially nitro-cellulose compounds, which are valuable in cases where glass is not suitable.

The mechanical and optical properties of such bodies have not, so far, been examined in very great detail, and the present paper describes some experimental evidence which has been obtained and which it is hoped to extend as opportunity occurs, since this has an important bearing on the study of stress problems arising in engineering practice.

The principal matters which are examined in the present communication are the mechanical properties of nitro-cellulose under pure tensile and bending stresses and the laws of its optical behaviour under these kinds of stress. In the course of the experimental study a considerable number of specimens have been examined, all of which are the manufacture of the British Xylonite Company.

The salient features of the material are its great flexibility and toughness, and the ease with which it can be drilled, turned or machined. By suitable adjustment of the condition of nitration of the body the hardness of the material can be varied through a considerable range, but owing to the difficulties created by the stress of war it has not been possible to make this investigation cover materials possessing a great range of hardness, and in fact all the specimens taken were originally selected

---

* L. N. G. FILON, "On the Variation with the Wave-length of the Double Refraction in Strained Glass," 'Camb. Phil. Soc. Proc.,' vol. XI., Part VI.; vol. XII., Part I.; and vol. XII., Part V.; see also 'Phil. Trans.,' A, vol. 207, and 'Roy. Soc. Proc.,' A, vols. 79 and 89. F. POCKELS, "Uber die Aenderung des Optischen Verhaltens verschiedener Gläser durch elastische Deformation," 'Annal. d. Physik,' 1902, and F. D. ADAMS and E. G. COKER, "The Cubical Compressibility of Rocks," 'Trans. Carnegie Institute.'

[Published October 18, 1920.

on account of their transparency and freedom from initial stress. The sheets from which the specimens are made differ greatly in age, one has been in stock for at least eight years, most of the others have been stored one or more years.

In order to examine the stress-strain properties of this material it is unnecessary to use a very delicate extensometer as the value of the modulus for direct stress is comparatively small, and for the purposes of these experiments a very simple form is employed consisting of a pair of clips attached respectively to a scale and a pointer, which latter slides over the scale and is kept in contact with it by suitable attachments. In order to examine the optical properties of the material while under stress, both scale and pointer are perforated to give a window opening, and thereby permit a beam of polarised light to be transmitted through the specimen under examination. With this instrument and special magnifying devices it is possible to estimate extensions of 0·0002 inch.

A preliminary examination of the problem set out above may be described with reference to some experiments on a specimen which was originally used in 1911 for determining the stresses in a notched tension member.*

Two bars, each 1 inch wide, were cut, at that time, from a clear plate of xylonite $\frac{3}{16}$ inch thick, and each was fashioned with notches of different sizes along the edges. One of these specimens has been used for the present test. A length of 6 inches was used for observations of the longitudinal strains, while the lateral strains required for determining POISSON'S ratio $\frac{1}{m} = \sigma$ have been measured by aid of a strain-measuring apparatus having a unit reading of $\frac{1}{1,000,000}$ inches. Unless otherwise stated these latter measurements are in the direction of the thickness of the material. As the details of the measuring apparatus and the cylindrical recorder used with it have already been described† they are not referred to further here.

Longitudinal Extension.—The specimen was examined in the polariscope under a moderate load and it was found that the stress was very uniformly distributed over a length of 6½ inches, but the remaining part of the parallel portion showed signs of unequal stress distribution owing to the enlarged ends. It was therefore marked off approximately into half-inch lengths over a total length of 6 inches, the exact distance being read to $\frac{1}{1000}$ inches.

YOUNG'S *Modulus.*—As it is convenient to start with a load of 20 lbs. on the specimen, a preliminary observation is made to determine the corresponding extension, and this value is allowed for in subsequent readings for convenience in plotting from a zero strain value.

---

* E. G. COKER, "The Effects of Holes and Semi-circular Notches on the Distribution of Stress in Tension Members," 'Proc. Phys. Soc.,' 1911.

† "Photo Elasticity for Engineers," by Prof. E. G. COKER, D.Sc., F.R.S., 'The Institute of Automobile Engineers,' November, 1917.

Six sets of observations are shown in fig. 1, the loads being applied in 20-lb. increments.

The maximum loads vary from 160 lbs. to 200 lbs., and after the strains reach their final value the load is again taken to its initial value and the scale read, to note the "semi-permanent set." The time interval between successive loads varies from one to two minutes. It is found that the stress-strain readings so obtained are approximately linear, except at the highest values, but in order to obtain as correct a value of Young's Modulus, E, as possible only measurements between 40 and 140 lbs. load are used, as the readings between these points are considered to be the most reliable. The strain corresponding to this difference of stress of 1256 lbs. per sq. inch is 0·00354 inch giving a value of E = 355,000.

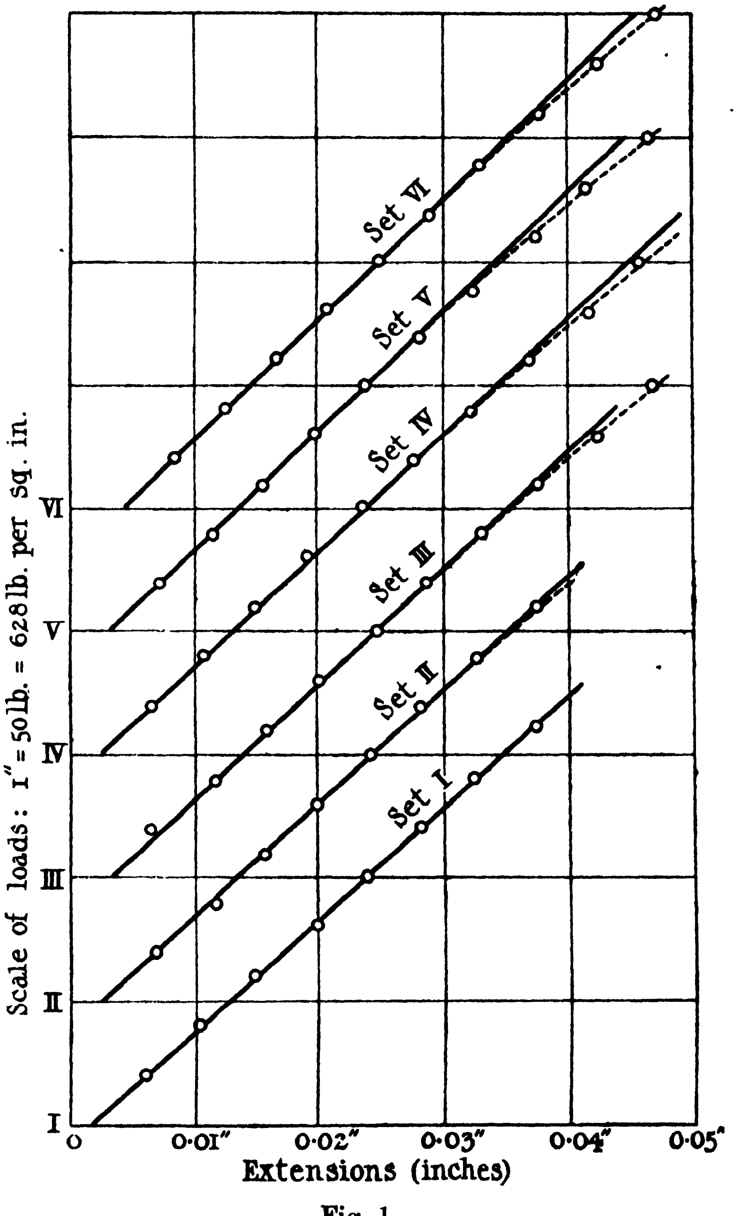


Fig. 1.

It will be observed that there is no very pronounced elastic limit, and that the curve is nearly straight up to 150 lbs. load (1900 lb./in.^{2}), which latter value may be taken as the elastic limit of the material. There is a "semi-permanent" set of 0·001 inches for each repetition of load, and a pronounced recovery between successive loadings especially with a short period of rest.

Measurements at Higher Stresses.—The spring balance used to measure the moderate loads in the above observations had a maximum capacity of 200 lbs., but for the higher stresses required, a balance recording up to 500 lbs. was necessary, the observations being made in a similar manner with readings on the magnified scale up to 400 lbs. load, and after this coarser readings were taken with the telescope. A maximum load of 476 lbs. (6000 lb./in.^{2}) was reached, but as the extension then increased very rapidly it was not possible to keep the load at this maximum value, moreover, as the stressing frame was of rather limited capacity for large strains, the test could not be carried to fracture, although a total extension of 1·211 inches was obtained. The observations also showed that the permanent extension was very uniform from section to section.

The condition of the material has in fact some resemblance to that of a mild steel which has been overstrained and allowed to rest. This is shown by a subsequent experiment in which the loading was repeated and the stress-strain properties examined anew. It was then found that the elastic limit of the material was still approximately at 150 lbs. load, corresponding to a stress of 1900 lbs. per sq. inch, but the modulus E had now risen to 502,000, measured in pound and inch units, as a result of the overstrain. The relations of load to extension for both conditions are shown in the accompanying fig. 2, but as the scales of load and extension are the same for both experiments the curve for overstrained material lies below that for unstrained material. It may be observed that the material possesses, in a marked degree, the property of contraction when the load is removed even when very much overstrained, and in this case when the full load of 300 lbs. was removed the semi-permanent extension was only 0·006 inch, and half of this disappeared with a few minutes rest.

Observations of lateral strain were also made with a suitable extensometer at several sections of the test bar, and their mean value for 100 lbs. load showed a strain of 0·00144, corresponding to a value of $m = \frac{1}{\sigma} = 2{\cdot}45$ where σ is POISSON's ratio.

The value of E is high as later experiments show, and this may possibly be due to an ageing effect, as in process of time the material appears to undergo some change, especially if the cut surfaces are not highly polished. This may probably be ascribed to the escape of a small portion of the volatile constituent of the material. It is also worthy of remark that the usual method of polishing appears to produce a thin outer layer which is harder than the interior, and this also has the effect of raising the value of E in thin specimens.

The effect of removing this thin layer of hard material has been under observation for some time, but in the experiments described here the flat sides are untouched, and the cut edges are unpolished although quite smooth.

Optical Properties.—There is considerable colour when an over-strained specimen is examined under no load in the polariscope. In the parallel part the colour is very uniform, and by comparing this with a previously unstrained specimen under load it appears that the permanent colour indicates a complex state of stress, since it could not be completely neutralised by a comparison tension piece, nor by bending the strained bar itself.

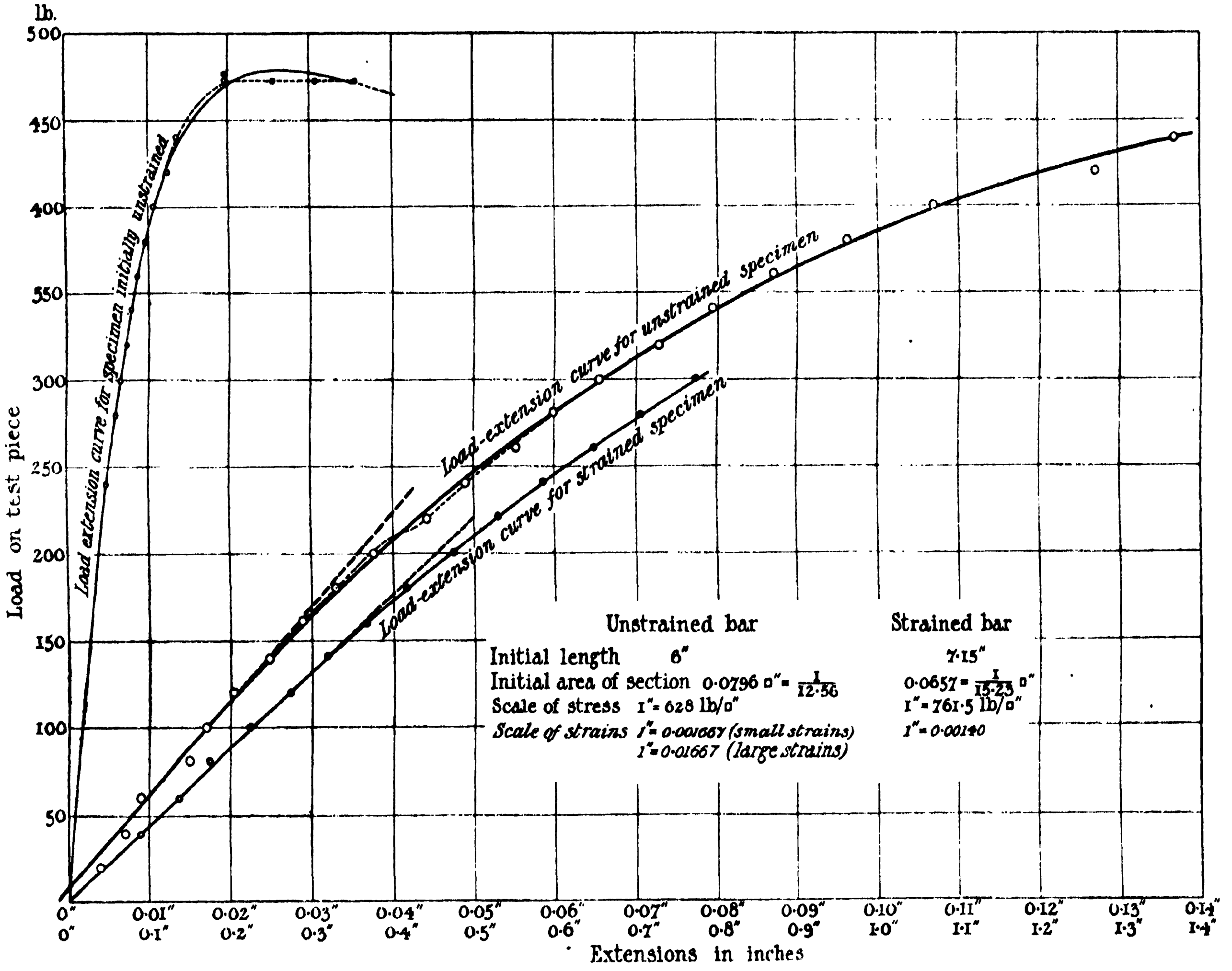


Fig. 2. Load-extension diagrams.

It may also be noted that nitro-cellulose with good optical properties is not apparently procurable, above ¼ inch thick, and it is difficult therefore to conform to the laws of similarity for the test specimens used in this investigation.

In later experiments the stress-strain properties of the material are examined both below and above the elastic limit, and the values of YOUNG'S modulus and POISSON'S ratio are measured for a number of specimens of different thickness and varying age. Especial attention is also directed to test the validity of the stress-optical law of this material since this is a matter of fundamental importance and little attention has so far been given to it.

The result of breaking the first three specimens showed that the extensometer arrangement was defective. The tiny indentations at the sides, where the extensometer was attached, so weakened the specimens that they all broke at one or other of these sections. The thinnest one naturally was affected most, so much so that it fractured with very little extension. This defect was partly corrected by cementing small fillets on to the specimen from which the extensometer clips were supported, but in spite of this some of the specimens fractured outside the gauge limits, showing that even in a ductile material the effects of enlarged ends prevents equalisation of stress near the change of section under any condition of load.

Stress Optical Determination.—It was originally intended to study the stress-optical properties of nitro-cellulose by analysing the light which traversed the material by means of a spectroscope ; but the necessary apparatus was rather difficult to procure, and it was convenient therefore to commence with a standard nitro-cellulose beam and use this for comparison with the optical phenomena observed in tension. The methods adopted here proved to be exceedingly well adapted for measurement of stress distribution beyond the elastic limit and are likely to be of great use hereafter. The comparison beam used is of rectangular section and is subjected to pure bending moment of known amount, and the stress at any point can therefore be calculated from the formula $f = \frac{\text{M}y}{\text{I}}$ without appreciable error.

It is generally assumed that the relative retardation of the polarised rays in a piece of optical material under moderate stress is proportional to the difference of principal stresses at the point, but this may not be correct and cannot be assumed to hold without experimental proof. Hence the stresses in the comparison beam are restricted to small values, so that the limit of proportionality of stress to strain is not passed in order to give an opportunity of examining the possibility of the law following a linear strain function or possibly some more complex variable. In order to make the retardation in such a beam sufficiently great to balance the retardation in the highly stressed specimens, the thickness of the beam should be large. This is most conveniently obtained by placing the several beams side by side, with their ends clamped and pinned together, as shown in fig. 3 in which several beams are so fastened together by plates A, to which extension levers B are also attached for supporting loads C depending from hangers D. This compound beam is supported on knife edges two inches apart, and when loaded has its central section sufficiently removed from the supports to give pure bending moment at the central section. The material of the beams is almost perfectly elastic up to and probably beyond 1600 lb./in.^{2}, but they are actually not stressed to more than 1300 lb./in.^{2}. In some cases as many as eight beams $\frac{1}{8}$ inch thick are used in this way, and a strong beam of light is then necessary to enable a comparison to be made with the tension member under observation. A carbon arc is then used as the source of light, but when only two or three thicknesses are employed the light from a Nernst lamp is sufficient, but in all cases the images are observed directly by eye instead of projecting on to a screen. The general arrangement of the apparatus is shown in the accompanying fig. 4, in which a plane polarised beam of white light from a NICOL'S prism A is transmitted through the tension specimen B, to which an extensometer C is secured, and is then focussed by a lens D on a horizontal slit in order that the light passing through the comparison beams shall be at the same level throughout. This thin pencil of light is again brought to parallelism before passing through the compound beam F and analyser G, and is finally focussed on a ruled glass slide H provided with an eye-piece J. The weight of the extension beams and hangers causes a bending moment in the beams which has been allowed for in all calculations of stress. In order to compare the different specimens one with another, an "equivalent stress" in each specimen is calculated, that is, such a stress as would produce the same relative retardation in a piece of nitro-cellulose of the same material as the standard beam, but of the thickness of the specimen under observation. Thus if the thickness of specimen is t and the stress in

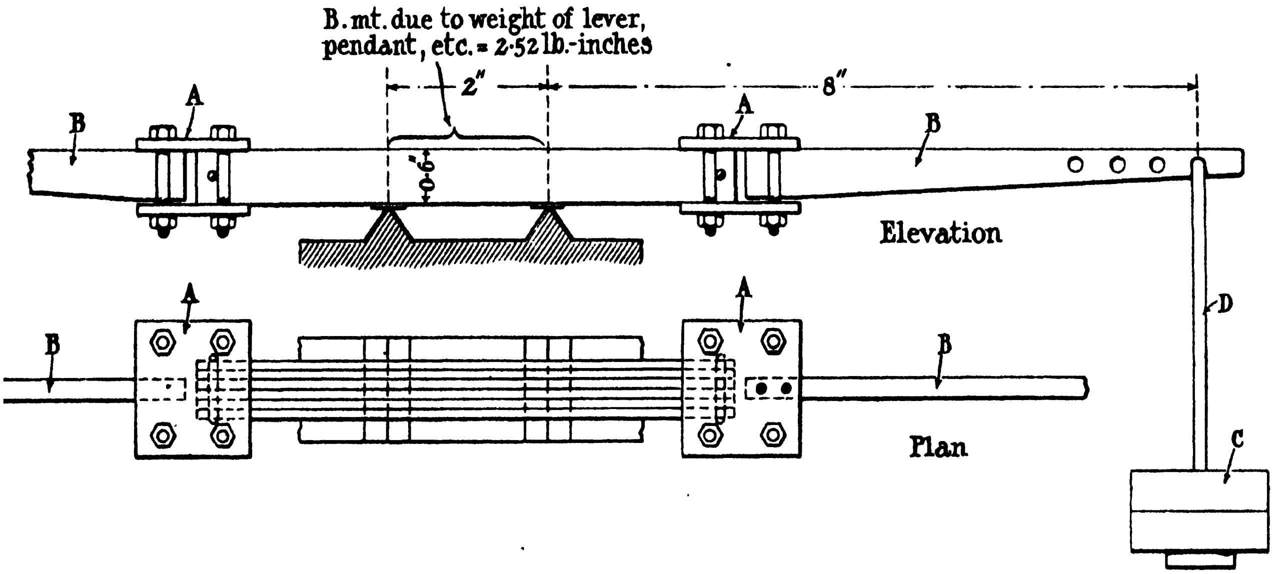

B.mt. due to weight of lever, pendant, etc. = 2·52 lb.-inches
2"
8"
A
B
0·6"
A
B
Elevation
A
B
A
B
D
Plan
C

Fig. 3. Beam Comparator.

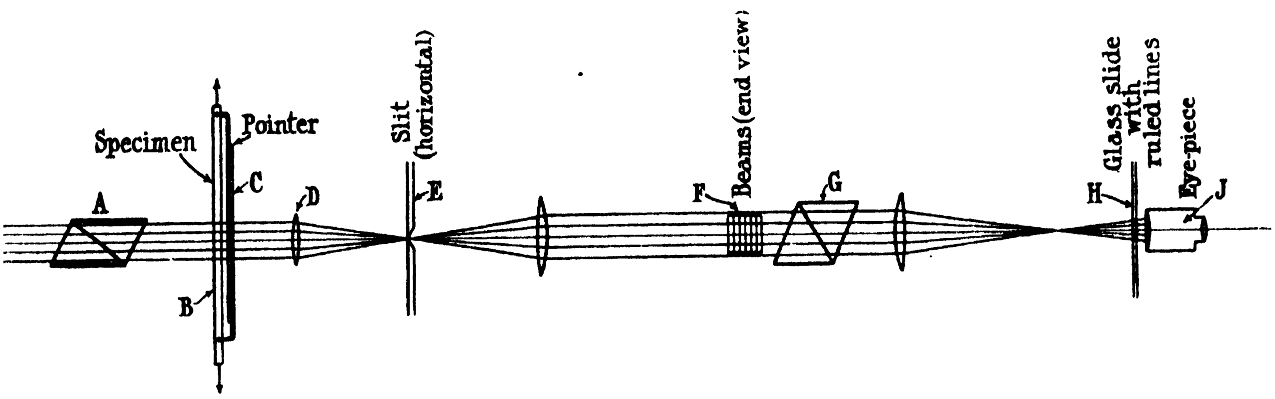


Fig. 4. Diagram of polariscope.

the beam at the points where the colour in the specimen is neutralised is f_0 and the corresponding thickness is t_0, we have the equivalent stress in the specimen $\frac{f_0 t_0}{t}$.

Now if M is the bending moment in the beam, d is its depth, and y is the distance from the neutral axis, then

$$f_0 = \frac{\mathrm{M}y}{\mathrm{I}} = \frac{\mathrm{M}y}{\frac{1}{12}t_0 d^3} = \frac{12\mathrm{M}y}{t_0 d^3},$$

so that the equivalent stress

$$f = \frac{f_0 t_0}{t} = \frac{12\mathrm{M}y}{td^3}.$$

Although as stated above the law of optical retardation is generally assumed to follow a linear law of stress difference, yet there is no apparent reason why it should not follow some other law, as for example a linear strain law or possibly contain terms involving squares of stress or strain. Some attempt has been made to find if the latter assumptions have any foundation, but if so the effects are within the limit of experimental error, and too small to be of any significance with the effect produced by a linear relation.

As regards the question whether this relation should be expressed in terms of stress or strain, it may be pointed out that an attempt is made here to test this with materials under direct stress, and that the validity of the law for combined stresses and strains still remains for consideration (apart from lateral strains, which are presumed to have no effect beyond altering the length of the path in which retardation takes place), but as in this case, if the standard, not stressed beyond the elastic limit is compared with another in which this condition is passed the experiments do in fact provide a means of discrimination, since in the standard, stress and strain are proportional, but are not so in general for the tension member. Hence if the form of the law of optical effect is assumed in terms of stress it does not exclude the possibility of finding from the experimental evidence whether it should not be expressed in terms of strain. We may, therefore, without loss of generality, take as an assumption the usual relation that relative retardation $\mathrm{R} = \mathrm{C}(\mathrm{P}-\mathrm{Q})\,t$ as a convenient expression where $(\mathrm{P}-\mathrm{Q})$ is the difference of principal stress $= f$, t is the thickness of the material and C is the stress-optical coefficient.

Let C_0 be the stress-optical coefficient of the standard beam. Then $\mathrm{R}_0 = \mathrm{C}_0 f_0 t_0$ for this beam.

When this latter is used to neutralise the retardation R in the specimen, since $\mathrm{R} = \mathrm{R}_0$, we have

$$\mathrm{C}ft = \mathrm{C}_0 f_0 t_0$$

But $t = t_0$, here and therefore

$$\mathrm{C}f = \mathrm{C}_0 f_0.$$

Now in general there is an initial retardation which is independent of any load. Let this correspond to stresses F, F_0. Then the condition $Cf = Cf_0$ becomes

$$C(f+F) = C_0(f_0+F_0)$$

or differentiating,

$$C \, . \, df = C_0 \, . \, df_0,$$

therefore

$$\frac{C}{C_0} = \frac{df_0}{df} = \frac{1}{\dfrac{df}{df_0}} = \text{reciprocal of slope of the stress/equivalent stress,}$$

which affords a convenient relation for examining the experimental data.

Turning now to the further experimental data upon the stress-strain properties of nitro-cellulose in tension a number of experiments have been made upon material of varying age and thickness, and these are plotted in the accompanying fig. 5, to show their characteristic properties under loads which sometimes exceed very considerably the elastic limits of the material.

With the thinnest specimens $\frac{1}{16}$ inch thick it was not found possible to obtain a reliable value of POISSON'S ratio, but YOUNG'S modulus, E, has been found, and the measurements plotted in fig. 5 show a characteristic feature that, although the first test is carried well beyond the elastic range, as soon as the load is removed a total extension of 0·0608 inch is reduced to 0·0070 inch or only 0·001 inch more than obtained at the commencement of this test. Moreover the value of the modulus changes less than 2 per cent. under these circumstances due to the earlier loading. It is also large as the skin effect is pronounced. These general characteristics are also observable in the measurements recorded in this figure for much thicker material if due allowance is made for the diminished effect of the surface layers. The capacity of returning to its original shape after high loads is still more marked in the next series of experiments on material $\frac{1}{8}$ inch thick, fig. 6, in which a stress of nearly 5000 lbs. per sq. inch is reached in the first experiment (curve 1) with nearly complete recovery, and when further loads with maxima varying from 4000 to 5000 lbs. per sq. inch are applied (curves 2 to 8) these give almost identical values of YOUNG'S modulus on the straight part of the curve until the ninth loading, where there is a sudden fall to E = 261,000 with an extension of 0·1045 inch corresponding to the initial load of 20 lbs. After this, with the considerable initial extension of 0·4290 inch, there is a great rise in the modulus. The value of POISSON'S ratio is very constant and is here found to reach the highest value of $m = \frac{1}{\sigma} = 2·7$.

Succeeding experiments on still thicker material confirm these results, and with the exception of the $\frac{3}{8}$-inch plate, the load extension curves agree in their linear character up to about 2000 lbs. per sq. inch, although the specimens differ in age and possibly also in composition. They have, however, the common feature of possessing

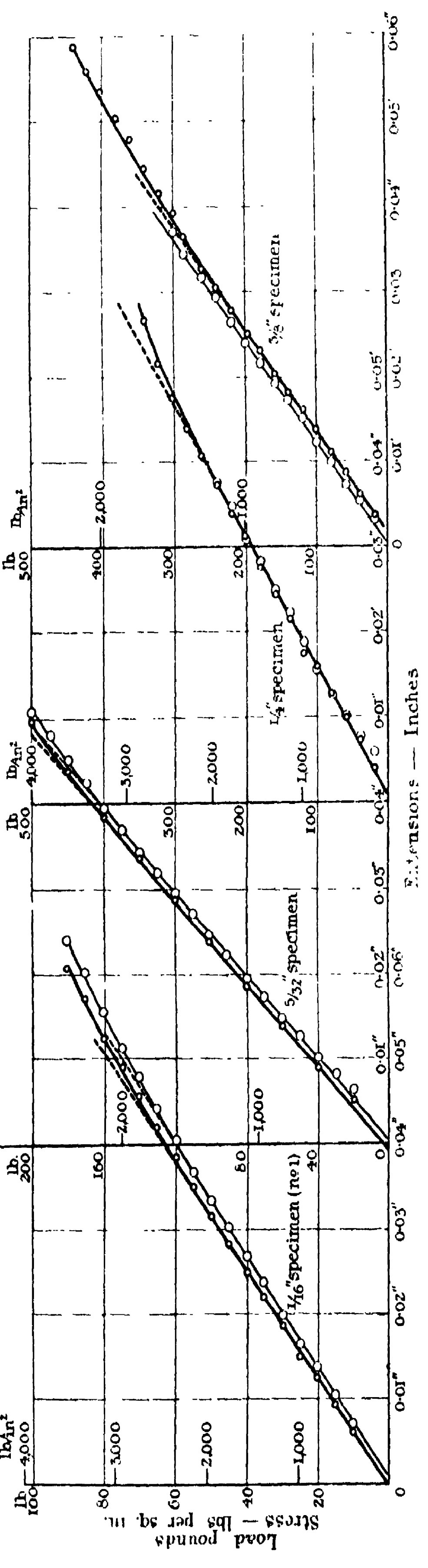


Fig. 5.

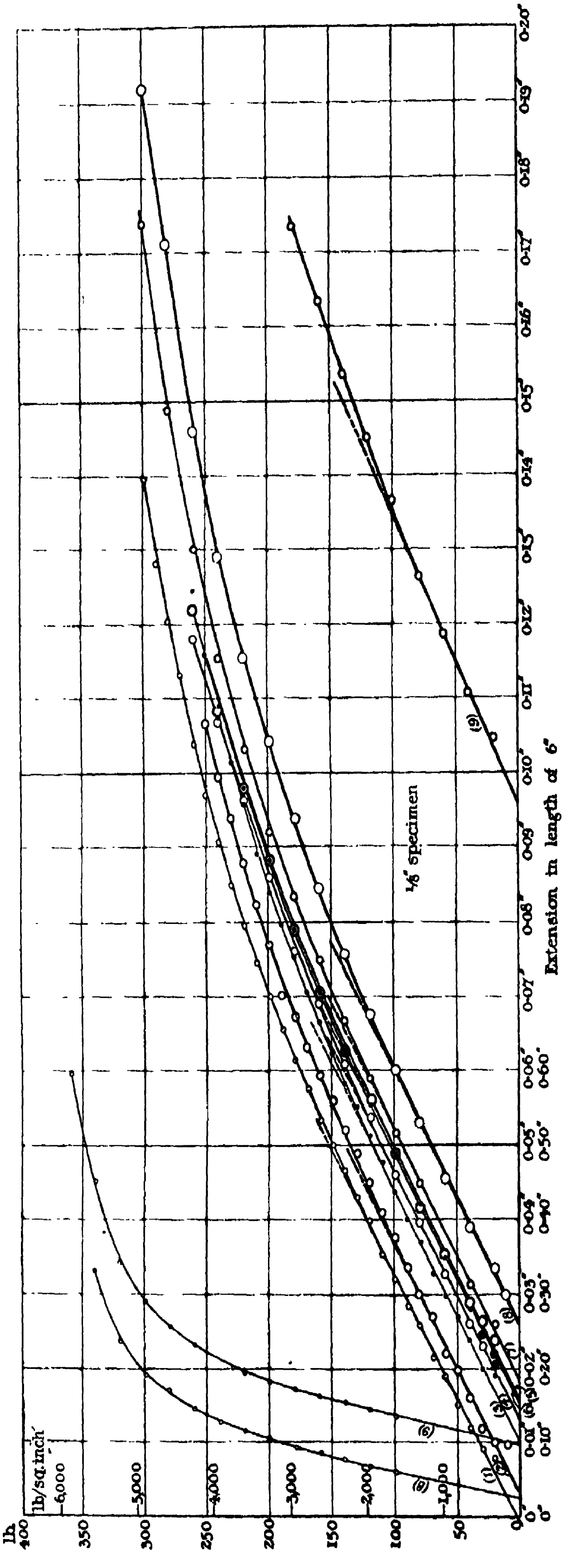


Fig. 6. Load-extension diagram—low stresses.

excellent optical properties and freedom from initial stress. The thickest plate, however, is exceptional, as its optical properties are poor.

Further experimental work on these materials is almost entirely devoted to the examination of the optical law of retardation under load, and for convenience all the data which follows is expressed as a stress or a strain, the units being pounds and inches, in which e is the strain under direct stress f, and the equivalent stress f_0 is obtained from the comparator beam. A typical example of these values is given in Table I., for material $\frac{1}{4}$ inch thick, as these measurements are referred to later for comparison with values of stress and strain obtained from spectrum observations.

TABLE I.

$\frac{1}{4}$-inch.			$\frac{1}{4}$-inch.		
e.	f.	f_0.	e.	f.	f_0.
0	0	73	0·0123	3315	3440
0·0003	160	220	0·0135	3505	3560
0·0007	319	322	0·0142	3665	3900
0·0013	478	440	0·0153	3825	4140
0·0020	637	660	0·0172	3980	4500
0·0027	797	880	0·0187	4140	4770
0·0033	956	1025	0·0208	4300	5100
0·0037	1115	1173	0·0212	4460	
0·0043	1274	1320	0·0237	4620	
0·0050	1433	1495	0·0287	4780	
0·0055	1594	1642	0·0320	1940	
0·0058	1752	1760	0·0370	5100	
0·0063	1912	1910	0·0553	5260	
0·0070	2070	2050	0·120	5420	
0·0077	2230	2200	0·157	5580	
0·0082	2390	2350	0·183	5740	
0·0087	2550	2540	0·247	5900	
0·0095	2710	2310	0·280	6055	
0·0102	2865	2860	0·340	6215	
0·0108	3025	3080	0·357	6375	
0·0117	3185	3230			

In the earlier experiment on the material $\frac{1}{16}$ inch thick, a fracture was obtained near the change of section and before the full extension developed, but still very nearly at the full load. It is included here (fig. 7) as, although the later parts of the stress-strain curves are not entirely satisfactory, this does not affect the problem in hand, since the stress-strain curve is not required very much beyond a pronounced yield in the material.

As a purely mechanical problem, however, there is a considerable amount of interest attaching to the accurate measurement of stress and strain over the whole of the plastic region, and it may be worth while at some future time to examine this with some care, especially if optical methods are applied to study the distribution of stress in purely plastic materials.

The stress-strain curves obtained for this thin material show a divergence from a linear law above 2000 lbs. per sq. inch whether plotted from the direct load or the optical stress measurements, but if the direct stress is plotted against its optical equivalent there is a definite linear law extending up to at least 4500 lbs. per sq. inch, and with only a small divergence at 5000 lbs. per sq. inch. The results in fact go to show that the law of retardation is linear as regards stress not only up to the elastic limit but actually to at least twice this range, where it is quite impossible for the strain to be linear. This result is shown in all the experiments on good optical material. Thus in plates $\frac{1}{8}$ inch thick where the elastic limit appears to be about 2250 lbs. per sq. inch, fig. 8, the corresponding value for f/f_0 shows no divergence from linearity until nearly double this amount, although the curve of f_0/e ceases to be linear at about 2500 pounds per sq. inch.

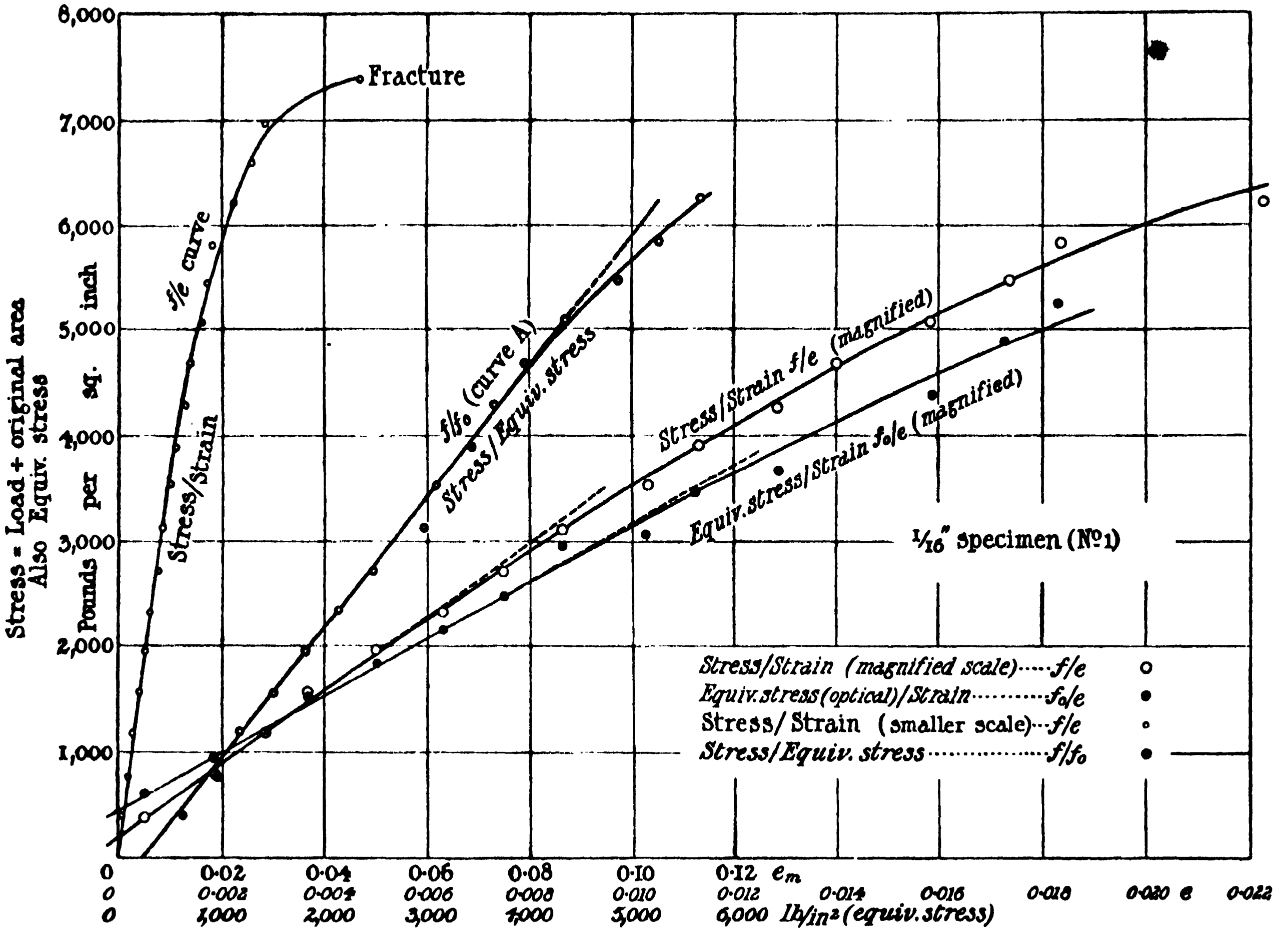


Fig. 7. Stress-strain curves of nitro-cellulose.

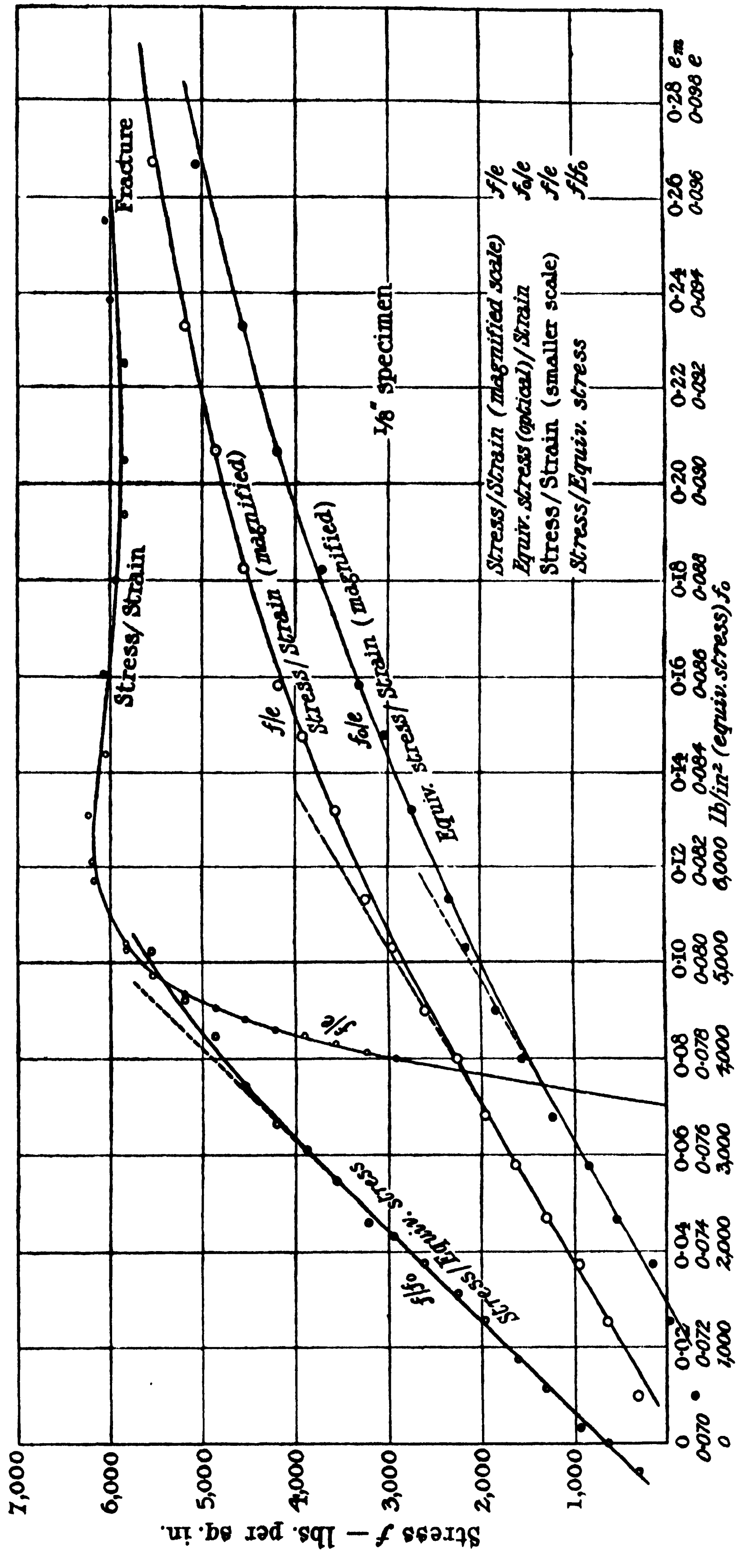

7,000
6,000
5,000
4,000
3,000
2,000
1,000
Stress f — lbs. per sq. in.
0 0·02 0·04 0·06 0·08 0·10 0·12 0·14 0·16 0·18 0·20 0·22 0·24 0·26 0·28 e_m
0·070 0·072 0·074 0·076 0·078 0·080 0·082 0·084 0·086 0·088 0·090 0·092 0·094 0·096 0·098 e
0 1,000 2,000 3,000 4,000 5,000 6,000 lb/in² (equiv. stress) f_0
Fracture
1/8" specimen
Stress/Strain
f/e Stress/Strain (magnified)
f_0/e Equiv. stress/Strain (magnified)
e/f
f/f_0 Stress/Equiv. stress
Stress/Strain (magnified scale) f/e
Equiv. stress (optical)/Strain f_0/e
Stress/Strain (smaller scale) f/e
Stress/Equiv. stress f/f_0

Fig. 8.

Somewhat similar results are obtained on plates $\frac{3}{16}$ inch thick, but in both experiments, fig. 9, the curves of f_0/e have a rather higher linear limit than the corresponding f/e curve, but here again the ratio f/f_0 is still linear to about the same range as in previous cases.

The case of plates $\frac{1}{4}$ inch thick, fig. 10, is more especially interesting from the fact that the stress-strain curve there shown is, at a later stage, obtained entirely from the optical effects observed from an analysis of the spectrum of a beam under uniform bending moment. It is sufficient to remark here that the f/f_0 curve shows a somewhat lower limit of linearity, although both the other curves have corresponding limits of 2000 lbs. per sq. inch.

When these curves are corrected for the change of cross-section which occurs as the test proceeds it is found, as fig. 10 shows, that the stress-strain curve f/e is perceptibly raised beyond the elastic limit and therefore tends more towards linearity, and the equivalent stress/strain curve is lowered and diverges still more from the linear relation. The stress/equivalent-stress curve has therefore a somewhat higher linear limit when this correction is made. Owing to the defective optical properties of still thicker material it was not found possible to examine these relations in a $\frac{3}{8}$-inch plate in a satisfactory manner.

Fracture.—The behaviour of nitro-cellulose at fracture is somewhat unusual for so ductile a material. As the load increases the section diminishes very uniformly at all parts removed from the enlarged ends, but there is little or no local contraction at any stage, and even at the fractured section, the cross-section differs but little from that at any other part of the bar, but after fracture there is a remarkable contraction in the total length accompanied by uniform expansion of the cross-section. This is shown in Table II., which gives a summary of the observations made and, except for one of the thin specimens and for the reasons given earlier, there is a recovery in length of from 6 to 9 per cent. after fracture. Various other measurements already described above are recorded here for convenient reference and also some ratios of the optical constants.

Spectrum Analysis of the Stress in a Beam.—The results of the optical examination appear to show the truth of the optical stress law for simple stress well beyond the elastic limit of the material, but the importance of this fundamental law makes it desirable to examine the matter in an independent way and possibly by a more rigid test than a comparison beam affords. An investigation of the optical phenomena presented by a beam under pure bending moment was made therefore on a rectangular strip $22\frac{1}{2}$ inches long, 1·005 inch deep and 0·2542 inch thick. Its specific gravity was approximately 1·361, this latter being determined at a temperature of 64° Fahr. by measurement of its volume and weighing in air.

The beam is supported as before on knife edges 2 inches apart, and the loading is applied at each end by dead weights having an overhang of $9\frac{3}{4}$ inches.

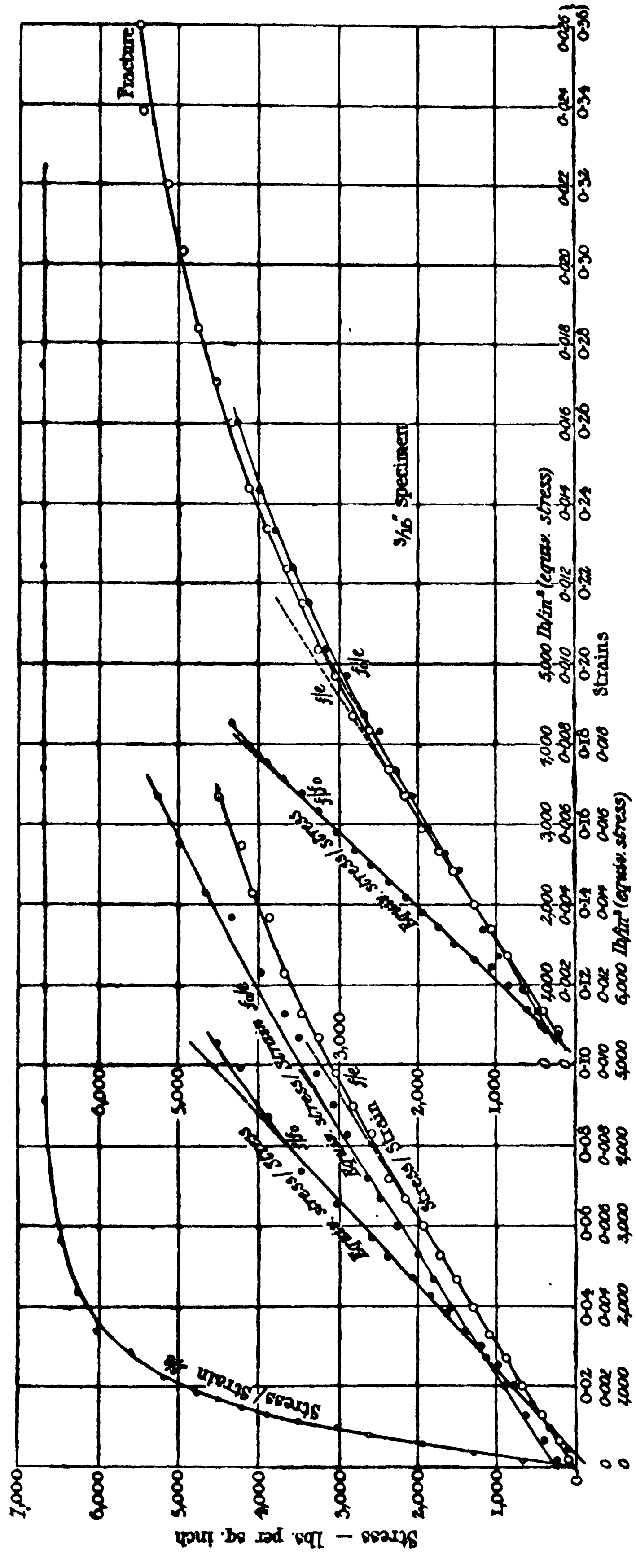

Fracture
3/16" specimen
Stress/Strain f/e
Birefr. stress/Strains
Stress/Strain
f/e
f₀/e
f/β₀
Birefr. stress/Strains
Strains
Stress — lbs. per sq. inch
0
1,000
2,000
3,000
4,000
5,000
6,000
7,000
0-02
0-04
0-06
0-08
0-10
0-12
0-14
0-16
0-18
0-20
0-22
0-24
0-26
0-28
0-30
0-32
0-34
0-36
lb/in² (equiv. stress)

Fig. 9.

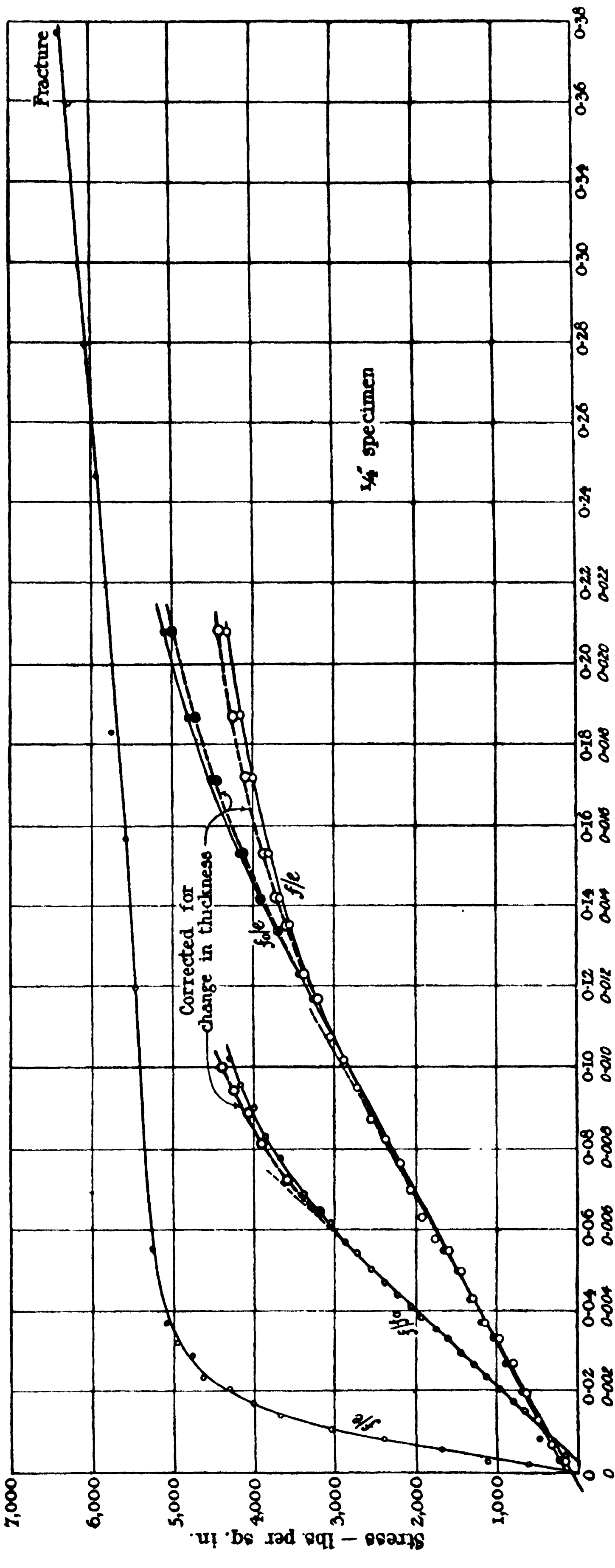

Fracture
1/4" specimen
Corrected for change in thickness
f/e
f₀/e
f/β₀
Stress — lbs. per sq. in.
0
1,000
2,000
3,000
4,000
5,000
6,000
7,000
0-02
0-04
0-06
0-08
0-10
0-12
0-14
0-16
0-18
0-20
0-22
0-24
0-26
0-28
0-30
0-32
0-34
0-36
0-38
lb/in²

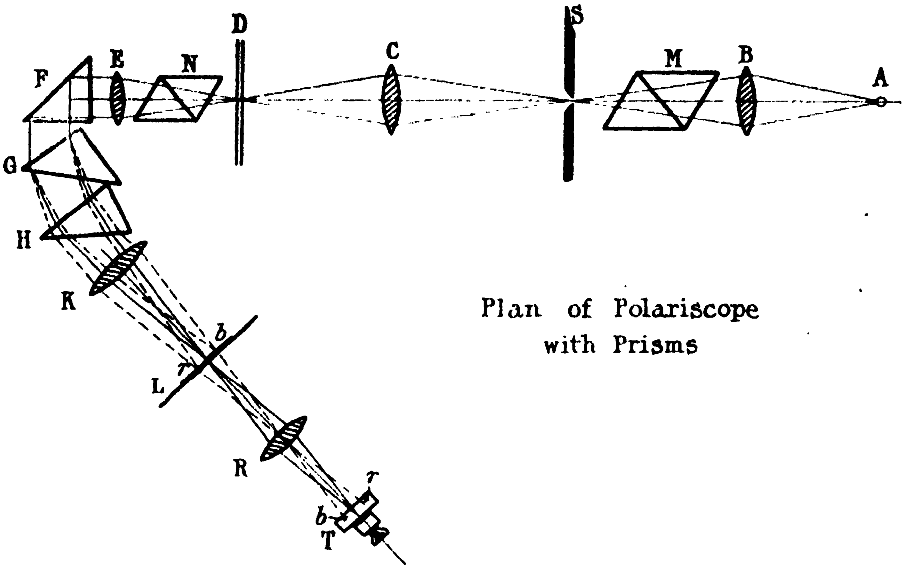
Fig. 10.

TABLE II.—Properties of the Specimens Tested.

Test length, 6-inch.

Specimen	$\frac{1}{16}''$ (No. 1)	$\frac{1}{16}'$ (No. 2)	$\frac{1}{8}''$	$\frac{5}{32}''$	$\frac{3}{16}''$	$\frac{1}{4}''$	$\frac{3}{8}''$
Width	0·5086″	0·5090″	0·5050″	0·4992″	0·4978″	0·4972″	0·4992″
Thickness	0·0505″	0·0506″	0·1213″	0·1517″	0·1863″	0·2533″	0·4045″
Area (in.^{2})	$0\cdot0257 = \frac{1}{38\cdot85}\square''$	$0\cdot0258 = \frac{1}{38\cdot75}\square''$	$0\cdot0613\square''$	$0\cdot07575 = \frac{1}{13\cdot2}\square''$	$0\cdot0928 = \frac{1}{10\cdot76}\square''$	$0\cdot1258 = \frac{1}{7\cdot97}\square''$	$0\cdot202 = \frac{1}{4\cdot96}\square''$
Modulus $\frac{\text{lb.}}{\text{in.}^2}$ . . .	362,000	355,000	295,000	324,000	313,000	309,000	251,000
POISSON's ratio . .	—	—	2·7	2·5	2·6	2·3	2·4
$\frac{C}{C_0}$ [optical] . . .	0·82	0·84	0·97	1·00	1·00	0·95	1·20 (?)
Maximum extension before fracture .	0·28″	1·66″	1·53″	2·14″	1·94″	2·14″	2·26″
Ultimate extension (after fracture) .	0	1·23″	1·07″	1·67″	1·54″	1·72″	1·74″
Recovery in length .	0·28″	0·43″	0·46″	0·47″	0·40″	0·42″	0·52″

The optical arrangements are modified for the new conditions as shown in the accompanying fig. 11. Light from the filament A of a Nernst lamp is focussed on to a vertical slit S by aid of a lens B, after passing through a NICOL's prism M and this narrow band is in turn focussed on the central section of the beam at D, and analysed by a second NICOL's prism N. A lens E placed at a convenient distance from the beam transmits this light as a parallel beam to a reflecting prism F, from which it passes through prisms G, H. The spectrum so obtained is focussed on a glass screen L ruled with lines $\frac{1}{100}$ inch apart, and provided with a micrometer eye-piece for measuring the ordinates of the bands observed.



Plan of Polariscope with Prisms

Fig. 11.

The field of view consists therefore of the spectrum of a Nernst lamp filament to which is added the effect produced by a narrow section of a beam of rectangular cross-section under pure bending moment. The relative retardation, owing to this latter stress effect, produces black bands in the field having a variable distance apart depending on the optical law of the retardation of the wave-length.

The general disposition of the field of view is shown in the accompanying fig. 12 in which bands of the first and second order appear on each side of the neutral axis C of the beam, and their co-ordinates are measured by reference to the graduations on the glass scale with the aid of a pair of parallel wires D, the positions of which can be adjusted vertically by a micrometer head E reading to $\frac{1}{6000}$ inch, while complete turns of the screw are obtained from a scale F on the left, which also appears in the field of view. In order to calibrate the horizontal scale the Nernst lamp and nitro-cellulose beam are removed, the Nicols rotated to parallelism, and a beam of solar light focussed on to the slit. The position of lines of known wave-length are noted with reference to the horizontal scale, and from these observations the constants in the equation

$$x_1 = A + \frac{B}{\lambda^2} + \frac{C}{\lambda^4}$$

are found for calibrating the positions of the black bands.

In the observations it is found that the depth of the beam does not appear quite constant throughout the field, an error due to the combined imperfections of the lenses and prisms employed. The maximum change of depth is about 2 per cent., and a correction is therefore necessary to reduce all vertical distances to a constant depth of beam.

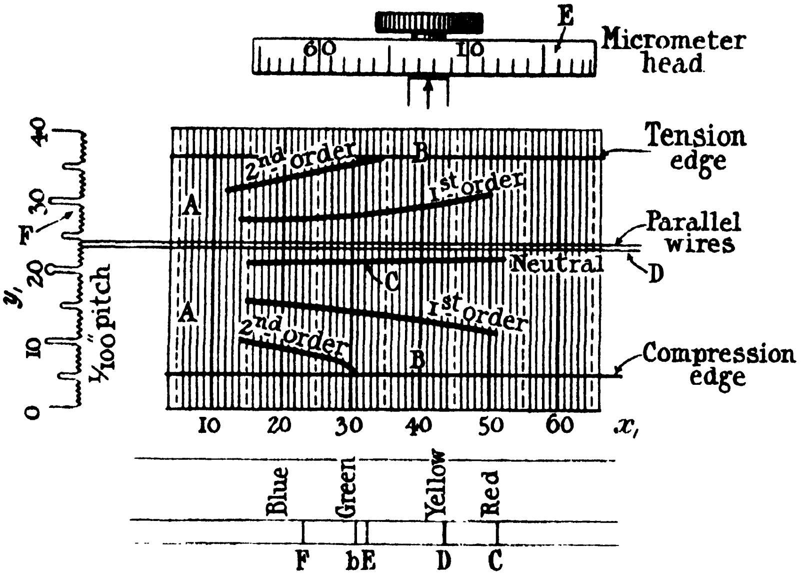


Fig. 12. View of spectrum.

The observations made are too numerous to give in detail, but typical examples of some measurements are shown in the accompanying fig. 13, in which the black bands, due to extinction of light, are drawn for a bending moment of 178·3 in pound and inch units. This, however, is not the exact appearance of the bands in the field of view owing to variation in the wave-length which alters the horizontal scale, but is here made uniform for plotting.

For some calculations, however, it is more convenient to show the form of the bands corresponding to a definite wave-length with a varying bending moment. Owing to the presence of a small amount of initial retardation in plates of nitro-cellulose, due to the method of manufacture, which leaves traces of initial stress, there is generally some slight difference between the bands on each side of the neutral axis, and a more accurate value is probable if the mean value for the two sides is taken.

If relative retardation is a linear function of the stress difference, these new abscissæ will represent the mean stress, but if the strain varies linearly they will also represent strains to another scale.

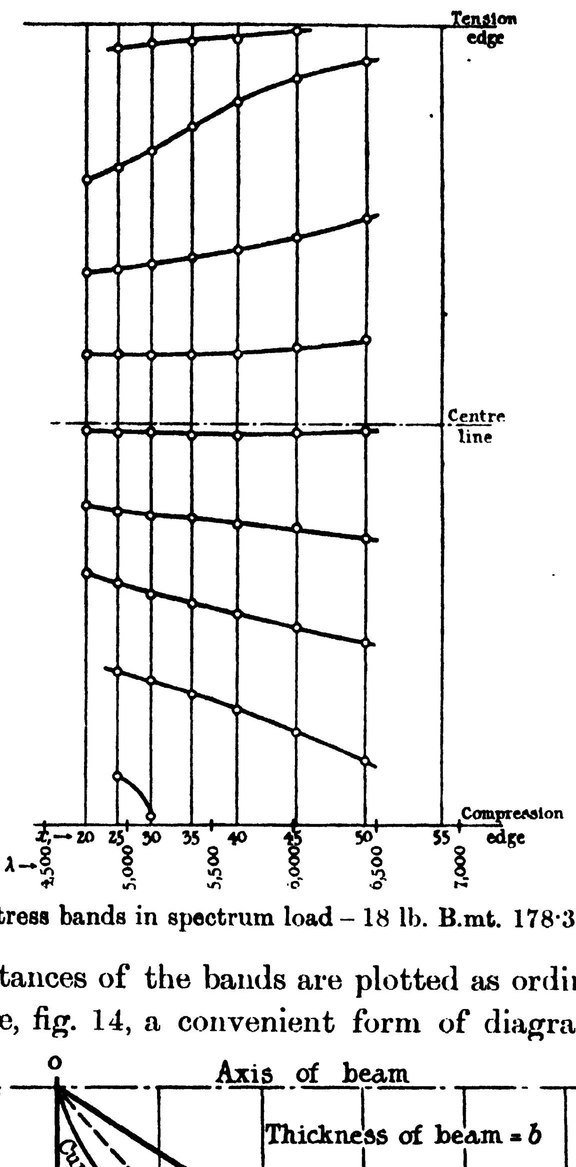


Fig. 13. Stress bands in spectrum load – 18 lb. B.mt. 178·3 lb.-inches.

If then the mean distances of the bands are plotted as ordinates against the order of the band as abscissæ, fig. 14, a convenient form of diagram OCF is obtained in

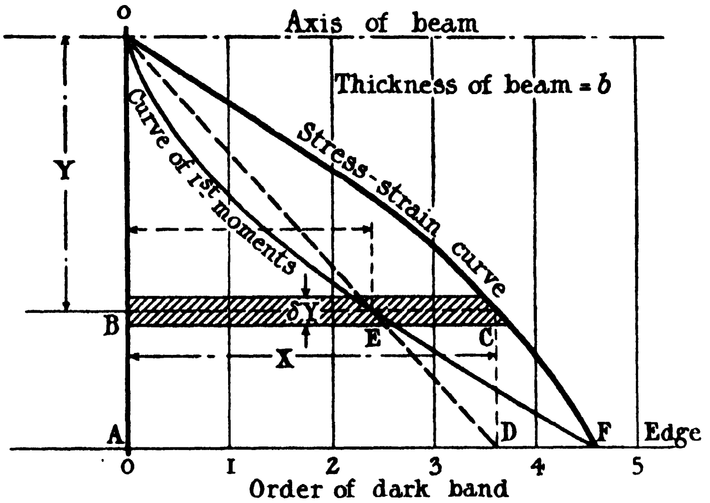


Fig. 14.

rectangular co-ordinates X, Y, in which Y is the distance from the neutral axis to a scale α and f is the stress to a scale β, or

$$y = Y \, . \, \alpha \, ; \qquad f = X \, . \, \beta, \text{ say.}$$

Now the bending moment

$$M = \int_{-\frac{1}{2}d}^{+\frac{1}{2}d} fb \,.\, y \,.\, dy$$

for a breadth $= b$ and a depth $= d$, or

$$M = b \,.\, \alpha^2 \,.\, \beta \int_{-\frac{1}{2}d}^{+\frac{1}{2}d} X \,.\, Y \,.\, dY$$

$= 2 \,.\, \alpha^2\beta \,.\, b \times$ first moment of the area of the diagram about the neutral axis.

If now any point C on the curve is projected on to the edge line at D and the line OD is drawn to the origin of co-ordinates intersecting the horizontal through C at E, then

$$BE : Y :: AD : AO = X : \frac{d}{2},$$

or

$$XY = BE \times \frac{d}{2}.$$

Hence $\int_{-\frac{1}{2}d}^{\frac{1}{2}d} X \,.\, Y \,.\, dY$ represents the area OEFAO $\times d$.

A typical example of one of these diagrams is shown in the accompanying fig. 15,

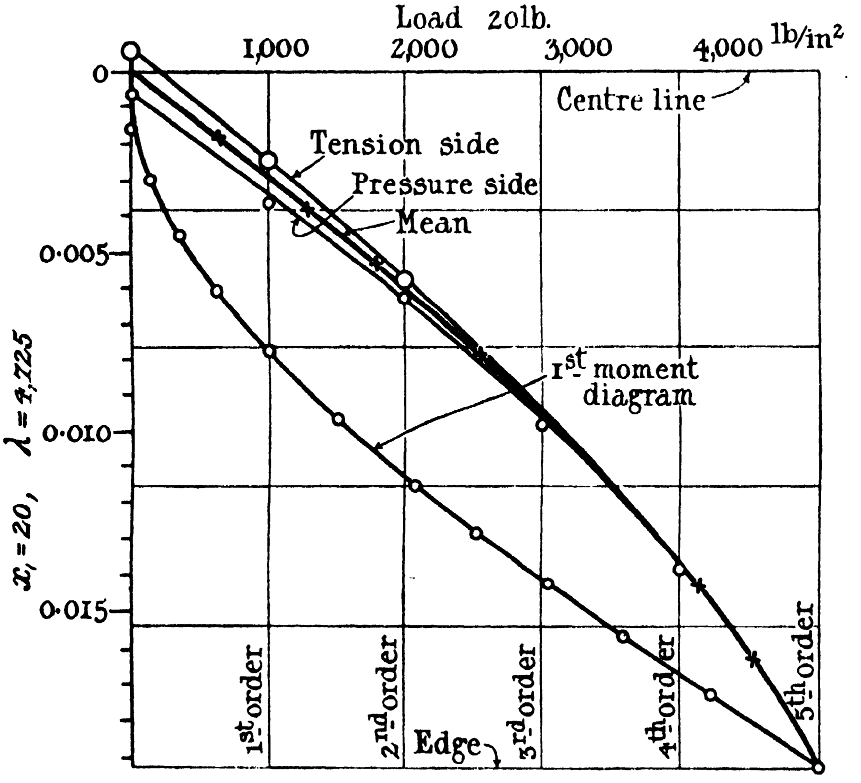


Fig. 15. Stress-strain curves.

of which about forty were actually prepared. The first moment areas M′, as determined by planimeter measurements, are shown in Table III., and when these are divided by the bending moment a value of $\alpha^2\beta b$ is obtained. If, however, the relative

TABLE III.—First Moments of Stress-strain Curves = M′.

x_1.	λ.	W = 20 lb.	18 lb.	16 lb.	14 lb.	12 lb.	10 lb.	9 lb.
		M = 198 lb./in.	178·3 lb./in.	158·7 lb./in.	139·2 lb./in.	119·6 lb./in.	100·0 lb./in.	90·2 lb./in.
20	4725	46·20	40·30	33·50	29·65	25·60	—	—
25	4925	43·15	38·40	32·15	28·65	24·35	—	—
30	5135	41·25	36·50	31·80	27·00	23·35	19·95	17·90
35	5380	39·85	36·40	31·10	25·75	21·20	18·25	—
40	5660	38·00	34·65	28·65	24·00	20·55	17·50	—
45	6015	36·65	31·90	27·85	23·00	19·25	17·00	—
50	6430	33·50	29·55	25·00	21·15	18·60	15·35	—

retardation is assumed to be independent of the wave-length, the mean value of $\frac{M'\lambda}{M}$ affords values of β corresponding to different wave-lengths. The values of β determined in this way are shown in the accompanying Table IV.

TABLE IV.

x_1.	λ.	β.
20	4725	920
25	4925	960
30	5135	1000
35	5380	1050
40	5660	1104
45	6015	1172
50	6430	1255

In order to determine the scale of strains the value of YOUNG's modulus E = 309,000, as taken from the measurements in tension within the elastic limit, is assumed to hold near the neutral axis of the beam for all loads, and since in this region we have the strain $e = Y . s$, where s is the scale for strains, then

$$E = \frac{df}{de} = \frac{\beta}{s} . \frac{dX}{dY},$$

or

$$s = \frac{\beta}{E} . \frac{dX}{dY}.$$

The slopes $\frac{dX}{dY}$ near the neutral axis are measured from the diagrams similar to those of fig. 15, and their values are shown in the accompanying Table V.

multiplied by the values of β appropriate to the wave-length. Their mean values afford measures of the strains as the table shows.

TABLE V.—Values of $\beta . \frac{\delta X}{\delta Y}$.

x_1 = 20	1165	1012	805	681	589	—	—
= 25	1167	1000	810	672	581	—	—
= 30	1180	1000	810	670	572	473	435
= 35	1180	998	840	657	559	466	—
= 40	1196	1032	848	657	563	468	—
= 45	1201	996	838	656	563	479	—
= 50	1205	997	831	677	564	467	—
Total . . .	8294	7035	5782	4670	3991	2353	435
Mean . . .	1185	1005	826	667	570	471	435 ?
Scale of strains $s = \frac{1}{E} . \beta . \frac{\delta X}{\delta Y}$.	$3 \cdot 84 \times 10^{-8}$	$3 \cdot 25 \times 10^{-8}$	$2 \cdot 67 \times 10^{-8}$	$2 \cdot 16 \times 10^{-8}$	$1 \cdot 85 \times 10^{-8}$	$1 \cdot 52 \times 10^{-8}$	$1 \cdot 39 \times 10^{-8}$
$\frac{1}{s} \times 10^{-8}$	0·260″	0·308″	0·375″	0·463″	0·541″	0·658″	0·720″

The data afforded by this method is therefore sufficient to construct a stress-strain curve entirely from these measurements of the bands due to retardation in the spectrum, and if the assumptions are correct, it ought to agree with a similar diagram constructed from data obtained independently. The stress-strain diagram

TABLE VI.—Stress-strain Values.

	x_1.	λ.	Strains.									
			0·002.	0·004.	0·006.	0·008.	0·010.	0·012.	0·014.	0·016.	0·018.	0·0192.
W = 20 lb.	20	4725	596	1210	1815	2350	2834	3290	3700	4090	4420	4570
	25	4925	612	1205	1827	2362	2830	3290	3672	4015	4340	4530
	30	5135	620	1230	1840	2360	2810	3250	3620	3950	4220	4340
	35	5380	629	1258	1855	2390	2850	3280	3644	3980	4270	4400
	40	5660	631	1262	1880	2424	2900	3330	3730	4020	4265	4360
	45	6015	647	1271	1860	2420	2910	3255	3765	4130	4450	4610
	50	6430	654	1295	1910	2440	2880	3295	3695	4000	4260	4370
Mean . . .			627	1247	1855	2394	2859	3284	3689	4026	4311	4454

obtained from spectrum observations gives the following values, of which Table VI. is a typical example, from which the mean values are obtained as follows:—

Strains	0·001	0·002	0·003	0·004	0·005	0·006	0·007	0·008
Stresses in lbs. per sq. inch	310	620	920	1240	1540	1845	2135	2400
Strains	0·009	0·010	0·011	0·012	0·013	0·014	0·018	0·0192
Stresses in lbs. per sq. inch	2665	2890	3055	3285	3460	3680	4310	4455

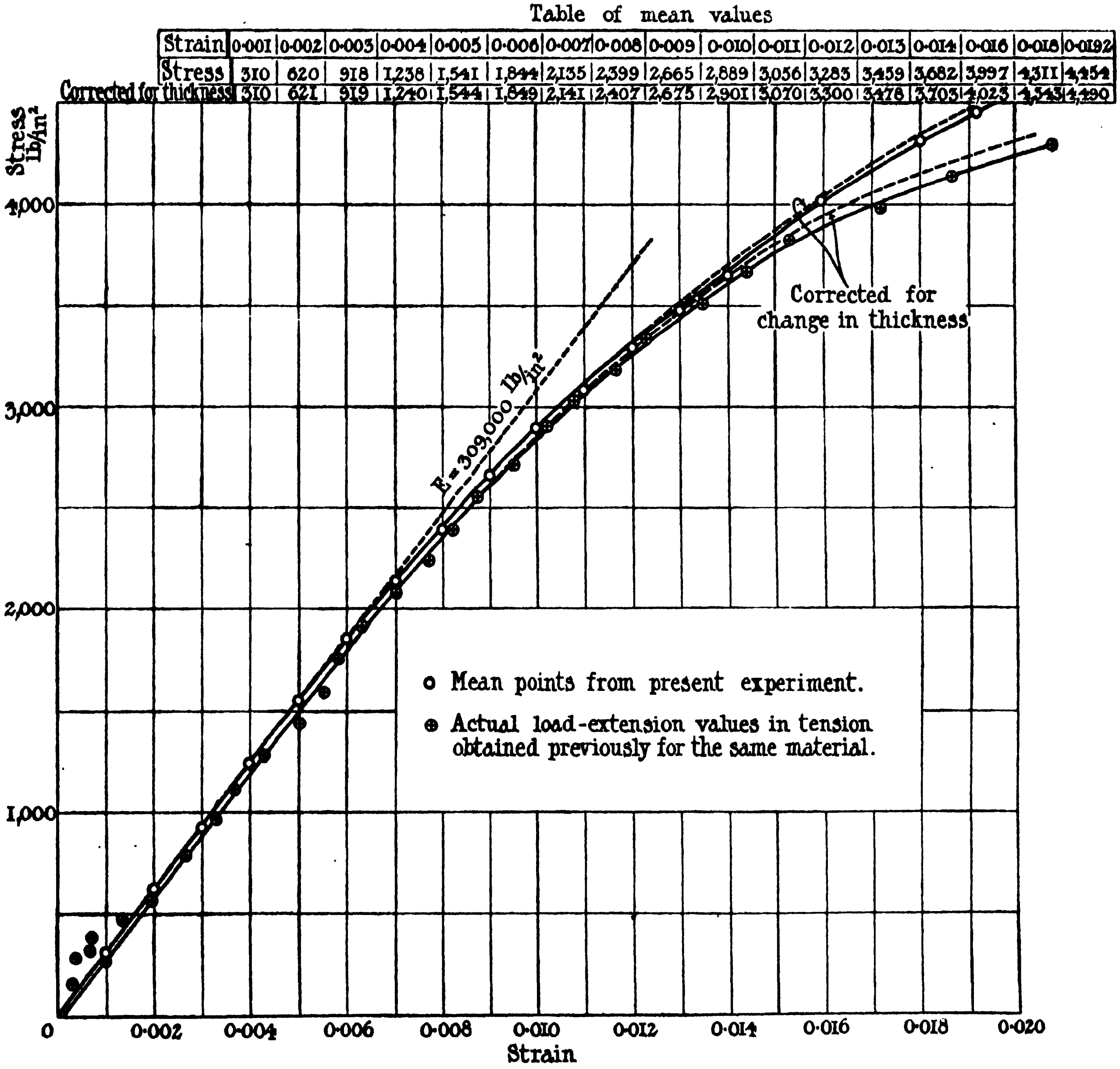


Table of mean values

Strain	0·001	0·002	0·003	0·004	0·005	0·006	0·007	0·008	0·009	0·010	0·011	0·012	0·013	0·014	0·016	0·018	0·0192
Stress	310	620	918	1,238	1,541	1,844	2,135	2,399	2,665	2,889	3,056	3,283	3,459	3,682	3,997	4,311	4,454
Corrected for thickness	310	621	919	1,240	1,544	1,849	2,141	2,407	2,675	2,901	3,070	3,300	3,478	3,703	4,023	4,543	4,490

Fig. 16. Comparison of stress-strain curves.

This information has already been obtained however by observations on precisely similar material under direct tension stress, Table II. and fig. 13, and on comparing plots of the two sets of data obtained, fig. 16, the agreement is seen to be a remarkably close one up to about 3500 lbs. per sq. inch. This agreement is improved if the changes in thickness are allowed for, since the corrected curves then lie closer together, and strengthen the evidence in favour of the law of optical retardation being an effect of stress and not of strain, and also that it is still a linear function much beyond the elastic limit of the material.

The whole of the evidence, in fact, appears to show that the transparent nitrocellulose examined obeys a linear stress optical law which holds up to approximately twice the range of the elastic limit of stress; and that within this range optical determinations of stress distribution may be relied upon.

In conclusion we desire to express our grateful thanks for the help afforded in this work by the Department of Scientific and Industrial Research, also for valuable suggestions from Prof. FILON, F.R.S., and Prof. PORTER, F.R.S., during its progress, and for the skilful assistance of Mr. F. H. WITHYCOMBE in preparing all the experimental apparatus required.

VI. *The Phenomena of Rupture and Flow in Solids.*

By A. A. GRIFFITH, *M. Eng.* (*of the Royal Aircraft Establishment*).

Communicated by G. I. TAYLOR, *F.R.S.*

Received February 11,—Read February 26, 1920.

CONTENTS.

		Page
1.	Introduction	163
2.	A Theoretical Criterion of Rupture	165
3.	Application of the Theory to a Cracked Plate	167
4.	Experimental Verification of the Theory	172
5.	Deductions from the Foregoing Results	177
6.	The Strength of Thin Fibres	180
7.	Molecular Theory of Strength Phenomena	183
8.	Extended Application of the Molecular Orientation Theory	186
9.	Practical Limitations of the Elastic Theory	192
10.	Methods of Increasing the Strength of Materials	194
11.	Application of the Theory to Liquids	195
12.	Summary of Conclusions	197

1. *Introduction.*

IN the course of an investigation of the effect of surface scratches on the mechanical strength of solids, some general conclusions were reached which appear to have a direct bearing on the problem of rupture, from an engineering standpoint, and also on the larger question of the nature of intermolecular cohesion.

The original object of the work, which was carried out at the Royal Aircraft Establishment, was the discovery of the effect of surface treatment—such as, for instance, filing, grinding or polishing—on the strength of metallic machine parts subjected to alternating or repeated loads. In the case of steel, and some other metals in common use, the results of fatigue tests indicated that the range of alternating stress which could be permanently sustained by the material was smaller than the range within which it was sensibly elastic, after being subjected to a great number of reversals. Hence it was inferred that the safe range of loading of a part, having a scratched or

[Published October 21, 1920.

grooved surface of a given type, should be capable of estimation with the help of one of the two hypotheses of rupture commonly used for solids which are elastic to fracture. According to these hypotheses rupture may be expected if (*a*) the maximum tensile stress, (*b*) the maximum extension, exceeds a certain critical value. Moreover, as the behaviour of the materials under consideration, within the safe range of alternating stress, shows very little departure from HOOKE's law, it was thought that the necessary stress and strain calculations could be performed by means of the mathematical theory of elasticity.

The stresses and strains due to typical scratches were calculated with the help of the mathematical work of Prof. C. E. INGLIS,* and the soap-film method of stress estimation developed by Mr. G. I. TAYLOR in collaboration with the present author.† The general conclusions were that the scratches ordinarily met with could increase the maximum stresses and strains from two to six times, according to their shape and the nature of the stresses, and that these maximum stresses and strains were to all intents and purposes independent of the absolute size of the scratches. Thus, on the maximum tension hypothesis, the weakening of, say, a shaft 1 inch in diameter, due to a scratch one ten-thousandth of an inch deep, should be almost exactly the same as that due to a groove of the same shape one-hundredth of an inch deep.

These conclusions are, of course, in direct conflict with the results of alternating stress tests. So far as the author is aware, the greatest weakening due to surface treatment, recorded in published work, is that given by J. B. KOMMERS,‡ who found that polished specimens showed an increased resistance over turned specimens of 45 to 50 per cent. The great majority of published results indicate a diminution in strength of less than 20 per cent. Moreover, it is certain that reducing the size of the scratches increases the strength.

To explain these discrepancies, but one alternative seemed open. Either the ordinary hypotheses of rupture could be at fault to the extent of 200 or 300 per cent., or the methods used to compute the stresses in the scratches were defective in a like degree.

The latter possibility was tested by direct experiment. A specimen of soft iron wire, about 0·028-inch diameter and 100 inches long, which had a remarkably definite elastic limit, was selected. This was scratched spirally (*i.e.*, the scratches made an angle of about 45 degrees with the axis) with carborundum cloth and oil. It was then normalised to remove initial stresses and subjected to a tensile load. Under these conditions the effect of the spiral scratches was to impart a twist to the wire, the twisting couple arising entirely from the stress-system due to the scratches. It was found that if the load exceeded a certain critical value, a part of the twist, amounting in some cases to 15 per cent., remained after the removal of the load. It was inferred that at this critical load the maximum stresses in the scratches reached the elastic limit of the material. This load was about one-quarter to one-third of that which caused the wire to yield as a whole, so that the scratches increased the maximum stress three or four times. The readings were quite definite even in the case of scratches produced by No. 0 cloth, which were found by micrographic examination to be but 10^{-4}-inch deep. Control experiments with longitudinal and circumferential scratches gave twists only 2 or 3 per cent. of those found with spiral scratches, and there was no permanent twist.

---

* "Stresses in a Plate due to the Presence of Cracks and Sharp Corners," 'Proc. Inst. Naval Architects,' March 14, 1913.

† 'Proc. Inst. Mech. Eng.' December 14, 1917, pp. 755-809.

‡ 'Intern. Assoc. for Testing Materials,' 1912, vol. 4A and vol. 4B.

This substantial confirmation of the estimated values of the stresses, even in very fine scratches, shows that the ordinary hypotheses of rupture, as usually interpreted, are inapplicable to the present phenomena. Apart altogether from the numerical discrepancy, the observed difference in fatigue strength as between small and large scratches presents a fundamental difficulty.

2. *A Theoretical Criterion of Rupture.*

In view of the inadequacy of the ordinary hypotheses, the problem of the rupture of elastic solids has been attacked from a new standpoint. According to the well-known "theorem of minimum energy," the equilibrium state of an elastic solid body, deformed by specified surface forces, is such that the potential energy of the whole system[*] is a minimum. The new criterion of rupture is obtained by adding to this theorem the statement that the equilibrium position, if equilibrium is possible, must be one in which rupture of the solid has occurred, if the system can pass from the unbroken to the broken condition by a process involving a continuous decrease in potential energy.

In order, however, to apply this extended theorem to the problem of finding the breaking loads of real solids, it is necessary to take account of the increase in potential energy which occurs in the formation of new surfaces in the interior of such solids. It is known that, in the formation of a crack in a body composed of molecules which attract one another, work must be done against the cohesive forces of the molecules on either side of the crack.[†] This work appears as potential surface energy, and if the width of the crack is greater than the very small distance called the "radius of molecular action," the energy per unit area is a constant of the material, namely, its surface tension.

In general, the surfaces of a small newly formed crack cannot be at a distance apart greater than the radius of molecular action. It follows that the extended theorem of minimum energy cannot be applied unless the law connecting surface energy with distance of separation is known.

---

[*] POYNTING and THOMSON, 'Properties of Matter,' ch. xv.

[†] The potential energy of the applied surface forces is, of course, included in the "potential energy, of the system."

There is, however, an important exception to this statement. If the body is such that a crack forms part of its surface in the unstrained state, it is not to be expected that the spreading of the crack, under a load sufficient to cause rupture, will result in any large change in the shape of its extremities. If, further, the crack is of such a size that its width is greater than the radius of molecular action at all points except very near its ends, it may be inferred that the increase of surface energy, due to the spreading of the crack, will be given with sufficient accuracy by the product of the increment of surface into the surface tension of the material.

The molecular attractions across such a crack must be small except very near its ends; it may therefore be said that the application of the mathematical theory of elasticity on the basis that the crack is assumed to be a traction-free surface, must give the stresses correctly at all points of the body, with the exception of those near the ends of the crack. In a sufficiently large crack the error in the strain energy so calculated must be negligible. Subject to the validity of the other assumptions involved, the strength of smaller cracks calculated on this basis must evidently be too low.

The calculation of the potential energy is facilitated by the use of a general theorem which may be stated thus: In an elastic solid body deformed by specified forces applied at its surface, the sum of the potential energy of the applied forces and the strain energy of the body is diminished or unaltered by the introduction of a crack whose surfaces are traction-free.

This theorem may be proved[*] as follows: It may be supposed, for the present purpose, that the crack is formed by the sudden annihilation of the tractions acting on its surface. At the instant following this operation, the strains, and therefore the potential energy under consideration, have their original values; but, in general, the new state is not one of equilibrium. If it is not a state of equilibrium, then, by the theorem of minimum energy, the potential energy is reduced by the attainment of equilibrium; if it is a state of equilibrium the energy does not change. Hence the theorem is proved.

Up to this point the theory is quite general, no assumption having been introduced regarding the isotropy or homogeneity of the substance, or the linearity of its stress-strain relations. It is necessary, of course, for the strains to be elastic. Further progress in detail, however, can only be made by introducing HOOKE's law.

If a body having linear stress-strain relations be deformed from the unstrained state to equilibrium by given (constant) surface forces, the potential energy of the latter is diminished by an amount equal to twice the strain energy.[†] It follows that the net reduction in potential energy is equal to the strain energy, and hence the total decrease in potential energy due to the formation of a crack is equal to the increase in strain energy less the increase in surface energy. The theorem proved above shows that the former quantity must be positive.

---

[*] The proof is due to Mr. C. WIGLEY, late of the Royal Aircraft Establishment.

[†] A. E. H. LOVE, 'Mathematical Theory of Elasticity,' 2nd ed., p. 170.

3. *Application of the Theory to a Cracked Plate.*

The necessary analysis may be performed in the case of a flat homogeneous isotropic plate of uniform thickness, containing a straight crack which passes normally through it, the plate being subjected to stresses applied in its plane at its outer edge.

If the plate is thin, its state is one of "plane stress," and in this case it may, without additional complexity, be subjected to any uniform stress normal to its surface, in addition to the edge tractions. If it is not thin, it may still be dealt with provided it is subjected to normal surface stresses so adjusted as to make the normal displacement zero. Here the plate is in a state of "plane strain." The equations to the two states are of the same form,* differing only in the value of the constants; they will therefore be taken together.

The strain energy may be found, with sufficient accuracy, in the general case where the edge-tractions are arbitrary; it is necessary in the present application, however, for the resulting stress-system to be symmetrical about the crack, as otherwise it is not obvious that the latter will remain straight as it spreads. The only stress distribution which will be considered, therefore, is that in which the principal stresses in the plane of the plate, at points far from the crack, are respectively parallel and perpendicular to the crack, and are the same at all such points. This is equivalent to saying that, in the absence of the crack, the plate would have been subjected to uniform principal stresses in and perpendicular to its plane. It is also necessary, on physical grounds, for the stress perpendicular to the crack and in the plane of the plate to be a tension, otherwise the surfaces of the crack are forced together instead of being separated, and they cannot remain free from traction.

In calculating the strain energy of the plate use will be made of the solution obtained by Prof. INGLIS for the stresses in a cracked plate, to which reference has already been made. The notation of Prof. INGLIS'S paper will be employed. In that notation α, β, are elliptic co-ordinates defined by the family of confocal ellipses; $\alpha = \text{const.}$ and the orthogonal family of hyperbolæ $\beta = \text{const.}$ The crack is represented by the limiting ellipse or focal line $\alpha = 0$. The axis of x coincides with the major axes, and the axis of y with the minor axes of the ellipses. The cartesian co-ordinates x, y, are connected with the elliptic co-ordinates α, β, by the relation

$$x + iy = c \cosh (\alpha + i\beta).$$

$R_{\alpha\alpha}$, u_{α}, are the tensile stress and displacement respectively along the normal to $\alpha = \text{const.}$

$R_{\beta\beta}$, u_{β}, are the corresponding quantities in the case of the normal to $\beta = \text{const.}$

$S_{\alpha\beta}$ is the shear stress in the directions of these normals.

c is the half-length of the focal line.

---

* A. E. H. LOVE, 'Mathematical Theory of Elasticity,' 2nd ed., p. 205.

h is the modulus of transformation, $\sqrt{\dfrac{2}{c^2(\cosh 2\alpha - \cos 2\beta)}}$.

μ is the modulus of rigidity of the material.

E is YOUNG'S modulus.

σ is POISSON'S ratio.

$p = 3 - 4\sigma$ in the case of plane strain, and

$\dfrac{3-\sigma}{1+\sigma}$ in the case of plane stress.

The state of uniform stress existing at points far from the crack (*i.e.* where α is large) will be specified by the three principal tensions P, Q and R. P is normal to the plate, and in the case of plane stress it is the same everywhere. Q and R are parallel respectively to the axes of x and y, and R is positive.

The strain energy of the plate is a quadratic function of P, Q and R, and hence, in accordance with the theorem proved above, the increase of strain energy due to the crack must be a positive quadratic function of P, Q and R. The general form of this function may be found by evaluating a sufficient number of particular cases.

The following particular cases are sufficient:—

I.—Q = R (and P = 0 in the case of plane stress).

Boundary of crack given by $\alpha = \alpha_0$.

The stresses are

$$R_{\alpha\alpha} = R\frac{\sinh 2\alpha\,(\cosh 2\alpha - \cosh 2\alpha_0)}{(\cosh 2\alpha - \cos 2\beta)^2}, \quad \dots \dots \dots \dots \quad (1)$$

$$R_{\beta\beta} = R\frac{\sinh 2\alpha\,(\cosh 2\alpha + \cosh 2\alpha_0 - 2\cos 2\beta)}{(\cosh 2\alpha - \cos 2\beta)^2}, \quad \dots \dots \quad (2)$$

$$S_{\alpha\beta} = R\frac{\sin 2\beta\,(\cosh 2\alpha - \cosh 2\alpha_0)}{(\cosh 2\alpha - \cos 2\beta)^2}, \quad \dots \dots \dots \dots \quad (3)$$

while the displacements are given by

$$\left.\begin{aligned} \frac{u_\alpha}{h} &= \frac{c^2 R}{8\mu}\{(p-1)\cosh 2\alpha - (p+1)\cos 2\beta + 2\cosh 2\alpha_0\} \\ \frac{u_\beta}{h} &= 0 \end{aligned}\right\} \quad \dots \quad (4)$$

The strain energy of the material within the ellipse α, per unit thickness of plate is

$$\tfrac{1}{2}\int_0^{2\pi}\frac{u_\alpha}{h}.R_{\alpha\alpha}.d\beta + \tfrac{1}{2}\int_0^{2\pi}\frac{u_\beta}{h}.S_{\alpha\beta}.d\beta. \quad \dots \dots \dots \quad (5)$$

On substituting and integrating, it is found that, as α becomes large, the strain energy tends towards the value

$$\frac{\pi c^2 R^2}{8\mu}\{\tfrac{1}{2}(p-1)e^{2\alpha}+(3-p)\cosh 2_{\alpha_0}\}. \quad \quad (6)$$

Hence W, the increase of strain energy due to the cavity α_0, is given by

$$W=\frac{\pi c^2 R^2}{8\mu}(3-p)\cosh 2\alpha_0 \quad \quad (7)$$

or, on proceeding to the limit, $\alpha_0 = 0$,

$$W=\frac{(3-p)\,\pi c^2 R^2}{8\mu} \quad \quad (8)$$

for a very narrow crack of length $2c$.

II.—$R = 0$ ($= P$ in the case of plane stress) $\alpha_0 = 0$.

Here the stresses are entirely unaltered by the crack, at every point of the plate except the two points $x = \pm c$, $y = 0$, where $R\alpha\alpha = -Q$. It follows that $W = 0$.

III.—$Q = R = 0$, $\alpha_0 = 0$.

Here, again, the stresses are unaltered, and $W = 0$.

The only positive quadratic function of P, Q and R which is compatible with these three particular cases is that given by equation (8); this is therefore the general form of W, and rupture is determined entirely by the stress R, perpendicular to the crack.

A point of some interest, with regard to equation (8), may be noticed in passing. Since W cannot be negative it follows that, in real substances, where μ is positive, $3 - p$ must be positive. Hence σ cannot be negative in real isotropic solids.

The potential energy of the surface of the crack, per unit thickness of the plate is

$$U = 4cT \quad \quad (9)$$

where T is the surface tension of the material.

Hence the total diminution of the potential energy of the system, due to the presence of the crack, is

$$W-U=\frac{(3-p)\,\pi c^2 R^2}{8\mu}-4cT. \quad \quad (10)$$

The condition that the crack may extend is

$$\frac{\partial}{\partial c}(W-U)=0,$$

or

$$(3-p)\,\pi c R^2 = 16\mu T, \quad \quad (11)$$

so that the breaking stress is

$$R = 2\sqrt{\frac{\mu T}{\pi\sigma c}}, \qquad (12)$$

in the case of plane strain, and

$$R = \sqrt{\frac{2ET}{\pi\sigma c}} \qquad (13)$$

in the case of plane stress.

Formula (13) has been verified experimentally. In connection with the experiments, interest attaches not only to the magnitude of R, but also to the value of the maximum tension in the material, which occurs at the extremities of the crack. This stress may be estimated if the radius of curvature of the boundary of the crack, at the points in question, can be found.

Expression (2) gives the maximum tension as

$$2R\sqrt{\frac{c}{\rho}} \qquad (14)$$

in case I. above, ρ being the radius of curvature at the corners of the elliptic crack. Prof. INGLIS shows that this expression may also be used, with little error, for cracks which are elliptic only near their ends. The foregoing expressions for the stresses are obtained, however, on the assumption that the displacements are everywhere so small that their squares may be neglected. At the corner of a very sharp crack, it cannot be assumed, without proof, that the change in ρ leaves formula (14) substantially unaffected.

In the case under consideration the displacements at the surface of the crack, due to a small tension dR at distant points, are given by

$$\left.\begin{aligned} \frac{u_\alpha}{h} &= \frac{c^2 dR}{E}(\cosh 2\alpha_0 - \cos 2\beta) \\ \frac{u_\beta}{h} &= 0 \end{aligned}\right\} \qquad (15)$$

Whence, by resolution, the displacements parallel respectively to the major and minor axes are

$$\left.\begin{aligned} u_x &= \frac{2dR}{E} c \sinh\alpha_0 \cos\beta \\ u_y &= \frac{2dR}{E} c \cosh\alpha_0 \sin\beta \end{aligned}\right\}, \qquad (16)$$

which may be written

$$\left.\begin{aligned} u_x &= \frac{2dR}{E} x \tanh\alpha_0 \\ u_y &= \frac{2dR}{E} y \coth\alpha_0 \end{aligned}\right\} \qquad (17)$$

Equations (17) show that the effect of the small stress $d\mathrm{R}$ on the elliptic cavity is to deform it into another ellipse. If a and b are the major and minor semi-axes of the ellipse, when the plate is subjected to a stress R, then, by (17),

$$\left.\begin{aligned}\frac{da}{d\mathrm{R}} &= \frac{2b}{\mathrm{E}}\\ \frac{db}{d\mathrm{R}} &= \frac{2a}{\mathrm{E}}\end{aligned}\right\} \quad \ldots\ldots \quad (18)$$

on making use of the relation $b = a \tanh \alpha_0$.

The solution of these simultaneous differential equations is

$$\left.\begin{aligned}a &= a_0 \cosh \frac{2\mathrm{R}}{\mathrm{E}} + b_0 \sinh \frac{2\mathrm{R}}{\mathrm{E}}\\ b &= a_0 \sinh \frac{2\mathrm{R}}{\mathrm{E}} + b_0 \cosh \frac{2\mathrm{R}}{\mathrm{E}}\end{aligned}\right\} \quad \ldots\ldots \quad (19)$$

where a_0 and b_0 are the values of a and b in the unstrained state.

With the help of equations (19) it is possible to find the maximum stress, F, due to an applied stress, R, taking account of the change in the shape of the cavity. From (2)

$$\frac{d\mathrm{F}}{d\mathrm{R}} = 2\frac{a}{b} \quad \ldots\ldots \quad (20)$$

whence

$$\mathrm{F} = 2\int_0^{\mathrm{R}} \frac{a_0 \cosh \frac{2\mathrm{R}}{\mathrm{E}} + b_0 \sinh \frac{2\mathrm{R}}{\mathrm{E}}\, d\mathrm{R}}{a_0 \sinh \frac{2\mathrm{R}}{\mathrm{E}} + b_0 \cosh \frac{2\mathrm{R}}{\mathrm{E}}}$$

$$= \mathrm{E} \log \left(\cosh \frac{2\mathrm{R}}{\mathrm{E}} + \frac{a_0}{b_0} \sinh \frac{2\mathrm{R}}{\mathrm{E}}\right), \quad \ldots\ldots \quad (21)$$

and in the case of a narrow crack which is elliptic only near its ends, $\frac{a_0}{b_0}$ may, as in (14), be replaced by $\sqrt{\frac{c}{\rho}}$.

In the general case, where Q is not equal to R, the quantity R — Q must be added to the value of F given by (21).

Formulæ (19) and (21) are not, of course, exactly true. The application of integration to equations (18) and (20) involves the assumption that the strains are so small that they can be superposed. If the strains are finite, this involves an error in the stresses depending on the square of the strains. In the case of ordinary solids, it is improbable that this assumption can alter the calculated stress by as much as 1 per cent.

4. *Experimental Verification of the Theory.*

In order to test formula (13), it was necessary to select an isotropic material which obeyed HOOKE's law somewhat closely at all stresses, and whose surface tension at ordinary temperatures could be estimated. For these reasons glass was preferred to the metals in common use. A comparatively hard English glass,* having the following properties, was employed:—

Composition—SiO_2, 69·2 per cent.; K_2O, 12·0 per cent.; Na_2O, 0·9 per cent.; Al_2O_3, 11·8 per cent.; CaO, 4·5 per cent.; MnO, 0·9 per cent.
Specific gravity—2·40.
YOUNG's modulus—$9·01 \times 10^6$ lbs. per sq. inch.
POISSON's ratio—0·251.
Tensile strength—24,900 lbs. per sq. inch.

The three last-named quantities were determined by the usual tension and torsion tests on round rods or fibres about 0·04-inch diameter and 3 inches long between the gauge points. The fibres had enlarged spherical ends which were fixed into holders with sealing wax. A slight load was applied while the wax was still soft, to ensure freedom from bending. The possible error of the extension measurements was about ± 0·3 per cent., and HOOKE's law was obeyed to this order of accuracy. No "elastic after-working" was observed with this glass, though more accurate measurements would doubtless have indicated its existence.

The problem of estimating the surface tension of glass, in the solid state, evidently requires special consideration. Direct determinations appeared to be impracticable, and ultimately an indirect method was decided on, in which the surface tension was found at a number of high temperatures and the value at ordinary temperatures deduced by extrapolation.

On the accepted theory of matter, intermolecular forces in solids and liquids consist mainly of two parts, namely, an attraction which increases rapidly as the distance between the molecules diminishes, balanced by a repulsion (the intrinsic pressure), which is due to the thermal vibrations of the molecules. It is reasonable to assume that the attraction, at constant volume, is sensibly independent of the temperature; this amounts merely to supposing that the attraction exerted by a molecule does not depend on its state of motion. On this view, the temperature variation, at constant volume, of the intermolecular forces is determined entirely by the change in thermal energy. Hence, it may be inferred, on the accepted theory of surface tension,† that the surface tension of a material, at constant volume, is equal to a constant diminished by a quantity proportional to the thermal energy of the substance. In the case of solids, nearly the same result should hold at constant pressure, as the temperature-volume change is small.

---

* Supplied in the form of test-tubes by Messrs. J. J. Griffin, Kingsway, London.

† POYNTING and THOMSON, 'Properties of Matter,' ch. xv.

The specific heat of glass is greater at high than at low temperatures, but the temperature coefficient is not large. Hence its surface tension may be expected to be nearly a linear function of the temperature, and extrapolation should be fairly reliable. This was found to be the case with the glass selected for the present experiments.

In the neighbourhood of 1100° C. the surface tension was found by QUINCKE's drop method. At lower temperatures this method was not satisfactory, on account of the large viscosity of the liquid glass; but between 730° C. and 900° C. the method described below was found to be practicable. Fibres of glass, about 2 inches long and from 0·002-inch to 0·01-inch diameter, with enlarged spherical ends, were prepared. These were supported horizontally in stout wire hooks and suitable weights were hung on their mid-points. The enlarged ends prevented any sagging except that due to extension of the fibres. The whole was placed in an electric resistance furnace maintained at the desired temperature. Under these conditions viscous stretching of the fibre occurred until the suspended weight was just balanced by the vertical components of the tension in the fibre. The latter was entirely due, in the steady state, to the surface tension of the glass, whose value could therefore be calculated from the observed sag of the fibre. In the experiments the angle of sag was observed through a window in the furnace by means of a telescope with a rotating cross wire. If w is the suspended weight, d the diameter of the fibre, T the surface tension, and θ the angle at the point of suspension between the two halves of the fibre, then, evidently,

$$\pi . d . \mathrm{T} . \sin \tfrac{1}{2}\theta = w.$$

For this method of determining the surface tension to be valid, it is evidently necessary that the angle of sag shall reach a steady value before the development of local contractions, arising from the instability of liquid cylinders, becomes appreciable. That this requirement is satisfied is shown by the following experimental results. After heating for two hours at about 750° C. the angle of sag of a particular fibre was 18°·25. Two hours later it had increased by less than 0°·1. The temperature was then raised momentarily to 940° C., and quickly reduced again to 750° C. The angle was then found to be 20°·2. After two hours further heating at 750° C. the angle had *decreased* to 18°·4, agreeing within permissible limits of error with the former value. That is to say, substantially the same limiting angle of sag was reached whether the initial angle was above or below that limit.

Above 900° C. it was found that the viscosity was insufficient to enable an observation to be made before the fibre commenced to break up into globules. Below 730° C., on the other hand, observations made on fibres of different diameters were inconsistent, the apparent surface tension being higher for the larger fibres. The obvious meaning of this result is that below 730° C. the glass used was not a perfect viscous liquid and hence the method was inapplicable. The transition from the viscous liquid state was quite gradual. The maximum tension (apart from surface tension) which could be permanently sustained, was zero at 730° C., 1·3 lbs. per sq. inch at 657° C., and 24 lbs.

per sq. inch at 540° C. At lower temperatures the rates of increase, both of this "solid stress" and the viscosity, were enormously greater. At 540° C. a fibre took about 70 hours to reach the steady state.

Table I. below gives the values of the surface tension obtained from these experiments. That for the temperature 1110° C. is the mean of five determinations by the drop method. The remaining figures were obtained from the sag of fibres.

TABLE I.—Surface Tension of Glass.

Temperature.	Surface Tension.
° C.	lb. per inch.
1110	0·00230
905	0·00239
896	0·00250
852	0·00249
833	0·00254
820	0·00249
801	0·00257
760	0·00255
745	0·00251
15	0·0031*

So far as they go, these figures confirm the deduction that the surface tension of glass is approximately a linear function of temperature. Moreover, as the actual variation is not great, the error involved in assuming such a law and extrapolating to 15° C. is doubtless fairly small. The value so obtained, 0·0031 lb. per inch, will be used in the present application.

Rigorously, expressions (13) and (21) above are true only for small cracks in large flat plates. In view, however, of the difficulties attendant on annealing and loading large flat glass plates, it was decided to perform the breaking tests on thin round tubes and spherical bulbs. These were cracked and then annealed and broken by internal pressure. The calculation cannot be exact for such bodies, but the error may obviously be reduced by increasing the ratio of the diameter of the bulb or tube to the length of the crack. It will be seen from the results of the tests that the variation of this ratio from two to ten caused little, if any, change in the bursting strength, and hence it may be inferred that the error in question is negligible for the present purpose.

The cracks were formed either with a glass-cutter's diamond, or by scratching with a hard steel edge and tapping gently. The subsequent annealing was performed by heating to 450° C. in a resistance furnace, maintaining that temperature for about one hour, and then allowing the whole to cool slowly. The question of the best annealing temperature required careful consideration, as it was evidently necessary to relieve the initial stresses due to cracking as much as possible, while at the same time keeping the temperature so low that appreciable deformation of the crack did not occur. It was found that the bursting strength increased with the annealing temperature up to about 400° C., while between 400° C. and 500° C. very little further change was perceptible. From this it was inferred that relief of the initial stresses was sufficient for the purpose in view at a temperature of 450° C.

---

* By extrapolation.

The principal stresses at rupture, Q and R, were calculated from the observed bursting pressure by means of the usual expressions for the stresses in thin hollow spheres and circular cylinders, the thickness of the glass near the crack being measured after bursting. In the case of the tubes the cracks were parallel to the generators, and provision was made for varying Q by the application of end loads.

In dealing with the longer cracks, leakage was prevented by covering the crack on the inside with celluloid jelly, the tube being burst before the jelly hardened. In the case of the smaller cracks leakage was imperceptible and this precaution was unnecessary.

The time of loading to rupture varied from 30 seconds to five minutes. No evidence was observed of any variation of bursting pressure with time of loading.

The results of the bursting tests are set down in Tables II. and III. below. $2c$ is the length of the crack, Q and R are the calculated principal stresses respectively parallel and perpendicular to the crack, and D is the diameter of the bulb or tube. The thickness of the bulbs was about 0·01 inch and the tubes 0·02 inch.

TABLE II.—Bursting Strength of Cracked Spherical Bulbs.

$2c$	D	Q	R	$R\sqrt{c}$.
inch.	inch.	lbs. per sq. inch.	lbs. per sq. inch.	
0·15	1·49	864	864	237
0·27	1·53	623	623	228
0·54	1·60	482	482	251
0·89	2·00	366	366	244

TABLE III.—Bursting Strength of Cracked Circular Tubes.

$2c$	D	Q	R	$R\sqrt{c}$.
inch.	inch.	lbs. per sq. inch.	lbs. per sq. inch.	
0·25	0·59	−621	678	240
0·32	0·71	−176	590	232
0·38	0·74	− 31	526	229
0·28	0·61	55	655	245
0·26	0·62	202	674	243
0·30	0·61	308	616	238

The average value of R $\sqrt{c}$ is 239, and the maximum 251.

According to the theory, fracture should not depend on Q, and R $\sqrt{c}$ should have, at fracture, the constant value

$$\sqrt{\frac{2ET}{\pi\sigma}}.$$

In the case of the glass used for these experiments, $E = 9 \cdot 01 \times 10^6$ lbs. per sq. inch, $T = 0 \cdot 0031$ lbs. per inch, and $\sigma = 0 \cdot 251$, so that the above quantity is equal to 266.

These conclusions are sufficiently well borne out by the experimental results, save that the maximum recorded value of R $\sqrt{c}$ is 6 per cent., and the average 10 per cent., below the theoretical value. It must be regarded as improbable that the error in the estimated surface tension is large enough to account for this difference, as this view would render necessary a somewhat unlikely deviation from the linear law.

A more probable explanation is to be obtained from an estimate of the maximum stress in the cracks. An upper limit to the magnitude of the radius of curvature at the ends of the cracks was obtained by inspection of the interference colours shown there. Near the ends a faint brownish tint was observed, and this gradually died out, as the end was approached, until finally nothing at all was visible. It was inferred that the width of the cracks at the ends was not greater than one-quarter of the shortest wave length of visible light, or about 4×10^{-6} inch. Hence ρ could not be greater than 2×10^{-6} inch.

Taking as an example the last bulb in Table II. and substituting in formula (21), it is found that

$$\frac{a_0}{b_0} = \sqrt{\frac{c}{\rho}} \geq 478$$

$$\frac{2R}{E} = 8 \cdot 13 \times 10^{-5},$$

whence the maximum stress $F \geq 344,000$ lbs. per sq. inch. The value given by the first order expression

$$F = 2R\sqrt{\frac{c}{\rho}}$$

is 350,000 lbs. per sq. inch.

A possible explanation of the discrepancy between theory and experiment is now evident. In the tension tests, the verification of HOOKE's law could only be carried to the breaking stress, 24,900 lbs. per sq. inch. There is no evidence whatever that the law is still applicable at stresses more than ten times as great. It is much more probable that there is a marked reduction in modulus at such stresses. But a decrease in modulus at any point of a body deformed by given surface tractions involves an increase in strain energy, and therefore in the foregoing experiments a decrease in strength. This is in agreement with the observations.

5. *Deductions from the Foregoing Results.*

The estimate of maximum stress obtained above appears to lead to a peremptory disproof of the hypothesis that the maximum tension in this glass, at rupture, is always equal to the breaking stress in ordinary tensile tests. If HOOKE's law was obeyed to rupture, and the squares of the strains were negligible, the maximum tension in the above cracked tube could not have been less than 344,000 lbs. per sq. inch; but, in the tensile tests, HOOKE's law was obeyed up to the breaking stress, the squares of the strains were negligible, and the maximum stress was only 24,900 lbs. per sq. inch. Hence the stresses could not have been the same in the two cases. Moreover, the order of the results obtained suggests (though this is not rigorously proved, as the assumptions have not been checked at stresses above 24,900 lbs. per sq. inch) that the actual strength may be more than ten times that given by the hypothesis.

Similar conclusions may be drawn regarding the " maximum extension," " maximum stress-difference " and " maximum shear strain " hypotheses which have been proposed from time to time for estimating the strength of brittle solids.

These conclusions suggest inquiries of the greatest interest. If the strength of this glass, as ordinarily interpreted, is not constant, on what does it depend ? What is the greatest possible strength, and can this strength be made available for technical purposes by appropriate treatment of the material ? Further, is the strength of other materials governed by similar considerations ?

Some indication of the probable maximum strength of this glass may be obtained from the bursting tests already described. There is no reason for supposing that, in those tests, the radii of curvature at the corners of the cracks were as great as 2×10^{-6} inch. It is much more likely that they were of the same order as the molecular dimensions. Considering, as before, the last bulb in Table II., and putting $\rho=2\times10^{-8}$ inch in formula (21), it is found that the maximum stress, F, is about 3×10^{6} lbs. per sq. inch. Elastic theory cannot, of course, be expected to apply with much accuracy to cases where the dimensions are molecular, on account of the replacement of summation by integration, and the probable diminution of modulus at very high stresses must involve a further error. Taking these circumstances into consideration, however, it may still be said that the probable maximum strength of the glass used in the foregoing experiments is of the order 10^{6} lbs. per sq. inch.

It is of interest to enquire at this stage whether there is any reason for ascribing similar maximum strengths to other materials. On the molecular theory of matter the tensile strength of an isotropic solid or liquid is of the same order as, though less than, its " intrinsic pressure," and may therefore be estimated either from a knowledge of the total heat required to vaporise the substance or by means of VAN DER WAAL's equation.* It may be noted that these methods of estimating the stress indicate that it should be, approximately at least, a constant of the material. TRAUBE* gives the following as the intrinsic pressures of a number of metals, at ordinary temperatures :—

---

* POYNTING and THOMSON, 'Properties of Matter,' ch. xv.

TABLE IV.—Intrinsic Pressures of Metals (TRAUBE).

Metal.	Intrinsic Pressure.
	lbs. per sq. inch.
Nickel	$4 \cdot 71 \times 10^6$
Iron	$4 \cdot 70 \times 10^6$
Copper	$3 \cdot 42 \times 10^6$
Silver	$2 \cdot 34 \times 10^6$
Antimony	$1 \cdot 74 \times 10^6$
Zinc	$1 \cdot 58 \times 10^6$
Tin	$1 \cdot 06 \times 10^6$
Lead	$0 \cdot 75 \times 10^6$

These are of the same order as the direct estimate obtained above for glass, but they are from 20 to 100 times the strengths found in ordinary tensile and other mechanical tests.

In the case of liquids, the discrepancy between intrinsic pressure and observed tensile strength is much greater. According to VAN DER WAAL'S equation, water has an intrinsic pressure of about 160,000 lbs. per sq. inch, whereas its tensile strength is found to be about 70 lbs. per sq. inch. It has been suggested that this divergence may be due to impurities, such as dissolved air, but DIXON and JOLY† have shown that dissolved air has no measurable effect on the tensile strength of water.

Thus the matters under discussion appear to be of general incidence, in that the strengths usually observed are but a small fraction of the strengths indicated by the molecular theory.

Some further discrepancies between theory and experiment may now be noticed. In the theory it is assumed that rupture occurs in a tensile test at the stress corresponding with the maximum resultant pull which can be exerted between the molecules of the material. On this basis the applied stress must have a maximum value at rupture, and hence, if intermolecular force is a continuous function of molecular spacing, the stress-strain diagram must have zero slope at that point. This, of course, is never observed in tensile tests of brittle materials ; in no case has any evidence been obtained of the existence of such a maximum anywhere near the breaking stress.

Again, the observed differences in strength as between static and alternating stress tests are at first sight inexplicable from the standpoint of the molecular theory, if the breaking load is regarded as the sum of the intermolecular attractions. According to the theory, large changes in the latter can only occur as a result of large changes in the thermal energy of the substance, such as would be immediately evident in alternating stress tests, if they took place.

---

* 'Zeitschr. für Anorganische Chemie,' 1903, vol. xxxiv., p. 413.

† 'Phil. Trans.,' B, 1895, p. 568.

Lastly, as indicated above, the strain energy at rupture of an elastic solid or liquid should on the molecular theory be of the same order as its heat of vaporisation. Hence rupture should be accompanied by phenomena, such as a large rise of temperature, indicative of the dissipation of an amount of energy of this order. It is well known that tensile tests of brittle materials show no such phenomena.

If, as is usually supposed, the materials concerned are substantially isotropic, there is but one hypothesis which is capable of reconciling all these apparently contradictory results. The theoretical deduction—that rupture of an isotropic elastic material always occurs at a certain maximum tension—is doubtless correct; but in ordinary tensile and other tests designed to secure uniform stress, the stress is actually far from uniform so that the average stress at rupture is much below the true strength of the material.

Now it may be shown, with the help of elastic theory, that the stress must be substantially uniform, in such tests, unless the material of the test-pieces is heterogeneous or discontinuous. It is known that all substances are in fact discontinuous, in that they are composed of molecules of finite size, and it may be asked whether this type of discontinuity is sufficient to account for the observed phenomena.

With the help of formula (13) above, this question may be answered in the negative. Formula (13) shows that a thin plate of glass, having in it the weakest possible crack of length $2c$ inch, will break at a tension, normal to the crack, of not less than $266/\sqrt{c}$ lbs. per sq. inch. This result, however, is subject to certain errors, and experiment shows that the true breaking stress is about $240/\sqrt{c}$ lbs. per sq. inch. But such a crack is the most extreme type, either of discontinuity or heterogeneity, which can exist in the material. Hence it is impossible to account for the observed strength, 24,900 lbs. per sq. inch, of the simple tension test specimens, unless they contain discontinuities at least $2 \times \left(\frac{240}{24,900}\right)^2$ inch, or say, 2×10^{-4} inch wide. This is of the order 10^{-4} times the molecular spacing.

The general conclusion may be drawn that the weakness of isotropic solids, as ordinarily met with, is due to the presence of discontinuities, or flaws, as they may be more correctly called, whose ruling dimensions are large compared with molecular distances. The effective strength of technical materials might be increased 10 or 20 times at least if these flaws could be eliminated.

It is easy to see why the presence of such small flaws can leave the strength of cracked plates, such as those of the foregoing experiments, practically unaffected. The most extreme case of weakening is that where there is a flaw very near the end of the crack and collinear with it. Here the result is merely to increase the effective length of the crack by less than 10^{-3} inch. This involves a weakening of less than 0·1 per cent.

6. *The Strength of Thin Fibres.*

Consideration of the consequences of the foregoing general deduction indicated that very small solids of given form, *e.g.*, wires or fibres, might be expected to be stronger than large ones, as there must in such cases be some additional restriction on the size of the flaws. In the limit, in fact, a fibre consisting of a single line of molecules must possess the theoretical molecular tensile strength. In this connection it is, of course, well known that fine wires are stronger than thick ones, but the present view suggests that in sufficiently fine wires the effect should be enormously greater than is observed in ordinary cases.

This conclusion has been verified experimentally for the glass used in the previous tests, strengths of the same order as the theoretical tenacity having been observed. Incidentally, information of interest has been obtained, somewhat unexpectedly, concerning the genesis of the flaws, and it has been found to be possible to prepare quite thick fibres in an unstable condition in which they have the theoretical strength.

Fibres of glass, about 2 inches long and of various diameters, were prepared. One end of a fibre was attached to a stout wire hanging on one arm of a balance, and the other end to a fixed point, the medium of attachment being sealing wax. A slight tension was applied while the wax was still soft, in order to eliminate bending of the fibre at the points of attachment. The other arm of the balance carried a beaker into which water was introduced from a pipette or burette. The weight of water necessary to break the fibre was observed, and the diameter of the latter at the fracture was found by means of a high-power measuring microscope. Hence the tensile strength was obtained.

At first the results were extremely irregular, though the general tendency of the strength to increase with diminishing diameter was clear. It was found that the irregularities were due to the dependence of the strength on the following factors :—

(1) The maximum temperature of the glass.—To secure the best results it was found necessary to heat the glass bead to about 1400° C. to 1500° C. before drawing the fibre.

(2) The temperature during drawing.—If the glass became too cool before drawing was complete, a weak fibre was obtained. This temperature could not be very closely defined, but it is perhaps the same as the limiting temperature of the viscous liquid phase, namely, 730° C. This effect made the drawing of very fine fibres a matter of some difficulty, as the cooling was so rapid.

(3) The presence of impurities and foreign bodies.

(4) The age of the fibre.—For a few seconds after preparation, the strength of a properly treated fibre, whatever its diameter, was found to be extremely high. Tensile strengths ranging from 220,000 to 900,000 lbs. per sq. inch were observed in fibres up to about 0·02 inch diameter. These strengths were estimated by measuring the radii to which it was necessary to bend the fibres in order to break them. They are therefore probably somewhat higher than the actual tenacities. The glass appeared to be

almost perfectly elastic up to these high stresses. The strength diminished, however, as time went on, until after the lapse of a few hours it reached a steady value whose magnitude depended on the diameter of the fibre.

Similar phenomena have been observed with other kinds of glass, and also with fused silica.

The relation between diameter and strength in the steady state was investigated in the following manner. Fibres of diameters ranging from 0·13 × 10^{-3} inch to 4·2 × 10^{-3} inch, and 6 inches long, were prepared by heating the glass to about 1400° C. to 1500° C. in an oxygen and coal-gas flame and drawing the fibre by hand as quickly as possible. The fibres were then put aside for about 40 hours, so that they might reach the steady state. The test specimens were prepared by breaking these fibres in tension several times until pieces about 0·5-inch long remained; these were then tested by the balance method already described. The object of this procedure was the elimination of weak places due to minute foreign bodies, local impurities and other causes.

Table V. below gives the results of these tests. Diameters are in thousandths of an inch, and breaking stresses in lbs. per sq. inch.

TABLE V.—Strength of Glass Fibres.

Diameter.	Breaking Stress.	Diameter.	Breaking Stress.
0·001 inch.	lbs. per sq. inch.	0·001 inch.	lbs. per sq. inch.
40·00	24,900*	0·95	117,000
4·20	42,300	0·75	134,000
2·78	50,800	0·70	164,000
2·25	64,100	0·60	185,000
2·00	79,600	0·56	154,000
1·85	88,500	0·50	195,000
1·75	82,600	0·38	232,000
1·40	85,200	0·26	332,000
1·32	99,500	0·165	498,000
1·15	88,700	0·130	491,000

It will be seen that the results are still somewhat irregular. No doubt more precise treatment of the fibres would lead to some improvement in this respect, but such refinement is scarcely necessary at the present stage.

The limiting tensile strength of a fibre of the smallest possible (molecular) diameter may be obtained approximately from the figures in Table V. by plotting reciprocals of the tensile strength and extrapolating to zero diameter. This maximum strength is found to be about 1·6 × 10^6 lbs. per sq. inch, which agrees sufficiently well with the rough estimate previously obtained from the cracked plate experiments.

---

* From the tensile tests previously described.

In 1858, Karmarsch* found that the tensile strength of metal wires could be represented within a few per cent. by an expression of the type

$$F = A + \frac{B}{d} \quad \ldots \ldots \ldots \ldots \quad (22)$$

where d is the diameter and A and B are constants. In this connection it is of interest to notice that the figures in Table V. are given within the limits of experimental error by the formula

$$F = 22{,}400 \frac{4.4 + d}{0.06 + d} \quad \ldots \ldots \ldots \ldots \quad (23)$$

where F is in lbs. per sq. inch and d is in thousandths of an inch. Within the range of diameters available to Karmarsch, this expression differs little from

$$F = 22{,}400 + \frac{98{,}600}{d}, \quad \ldots \ldots \ldots \ldots \quad (24)$$

which is of the form given by Karmarsch. Moreover, the values of B found by him for the weaker metals, *e.g.*, silver and gold, in the annealed state, are of the same order as that given by formula (24) for glass.

To a certain extent this correspondence suggests that the mechanism of rupture, as distinct from plastic flow, in metals, is essentially similar to that in brittle amorphous solids such as glass.

The remarkable properties of the unstable strong state referred to on p. 180 above are exhibited most readily in the case of clear fused silica. If a portion of a rod about 5 mm. diameter be made as hot as possible in an oxyhydrogen flame, and then drawn down to, say, 1 mm. or less and allowed to cool, the drawn-down portion may be bent to a radius of 4 or 5 mm. without breaking, and if then released will spring back almost exactly to its initial form. If instead of being released it is held in the bent form it will break spontaneously after a time which usually varies from a few seconds to a few minutes, according to the degree of flexure. To secure the best results the drawing should be performed somewhat slowly.

When fracture occurs it is accompanied by phenomena altogether different from those associated with the fracture of the normal substance. The report is much louder and deeper than the sharp crack which accompanies the rupture of an ordinary silica or glass rod, and the specimen is invariably shattered into a number of pieces, parts being frequently reduced to powder. This shattering is not confined to the highly stressed drawn-out fibre; it usually occurs also in the unchanged parts of the original thick rod and sometimes in the free ends, which are not subjected to the bending moment. The experiment is most striking, for it appears at first sight that the 5 mm. rod has been broken by a couple applied through the fibre, which may be only 0·5 mm. in diameter. As a matter of fact, however, the shattering is probably merely one of the means of dissipating the strain energy of the strong fibre, which at fracture is perhaps 10,000 times that of silica in the ordinary weak state. An elastic wave is doubtless propagated from the original fracture, and the stresses due to this wave shatter the rod.

---

* ‘Mittheilungen des gew. Ver. für Hannover,’ 1858, pp. 138–155.

Confirmation of this view is obtainable if the fibre is broken by twisting instead of by bending. The thick part of the rod is in this case found to contain a number of spiral cracks, at an angle of about 45° to the axis, showing that the material has broken in tension, but the cracks run in both right- and left-handed spirals, so that the surface of the rod is divided up into little squares. This shows that the cracking must be due to an alternating stress, such as would result from the propagation of a torsional wave along the rod.

Another phenomenon which has been observed in these fibres is that fracture at any point appears to cause a sudden large reduction in the strength of the remaining pieces. Thus, in one case a glass fibre was found to break in bending at an estimated stress of 220,000 lbs. per sq. inch. One of the pieces, on being tested immediately afterwards, broke at about 67,000 lbs. per sq. inch.

7. *Molecular Theory of Strength Phenomena.*

From the engineering standpoint the chief interest of the foregoing work centres round the suggestion that enormous improvement is possible in the properties of structural materials. Of secondary, but still considerable, importance is the demonstration that the methods of strength estimation in common use may lead in some cases to serious error.

Questions relating to methods of securing the indicated increase in tenacity, or of eliminating the uncertainty in strength calculations, can scarcely be answered without some more or less definite knowledge of the way in which the properties of molecules enter into the phenomena under consideration. In this connection it is of interest to enquire whether any indication can be obtained of the nature of the properties which are requisite for an explanation of the observed facts.

For this purpose it is convenient to start with molecules of the classical type, whose properties may be defined as (*a*) a central attraction between each pair of molecules which decreases rapidly as their central distance increases, and which depends only on that distance and the nature of the molecules ; (*b*) translational and possibly rotational vibrations whose energy is the thermal energy of the substance. In the unstrained state, the kinetic reactions due to (*b*) balance the central attractions (*a*).

In a body composed of such molecules, the flaws which have been shown to exist in real substances might consist of actual cracks. But experiment shows that under certain conditions the strength of glass diminishes with lapse of time. On the present hypothesis this would require the potential energy of the system to increase spontaneously by the amount of the surface energy of the cracks. This view must therefore be regarded as untenable.

Again, the observed weakening might conceivably occur if at any instant the vibrations of a large number (at least 10^8) of near molecules synchronised and were in phase, provided the energy of these molecules was approximately that corresponding with the temperature of ebullition of the substance. Except in the case of a material very near its boiling point, the probability of such an occurrence must be so small as to be quite negligible. Hence this hypothesis also must be discarded.

The foregoing discussion seems to suggest that the assumed type of molecule is too simple to permit of the construction of an adequate theory. An increase in generality may be obtained by supposing that the attraction between a pair of molecules depends not only on their distance apart, but also on their relative orientation. The properties of crystals seem definitely to require the molecules of anisotropic materials to be of this type, but those of isotropic substances have usually been assumed to be of the simpler kind. In view of the author, however, molecular attraction must be a function of orientation even in substances, such as glass, metals and water, which are usually referred to as "isotropic."

Consider a solid made up of a number of such molecules, initially oriented at random. Doubtless the mechanical properties of the substance, while it is in this amorphous condition, will differ little from those of a substance composed of molecules of the simpler type, having an attraction of appropriate strength. If this is so, the tensile strength of the material must be that corresponding with its average intrinsic pressure.

In general, however, this initial condition cannot be one of minimum potential energy.

It is clear that under suitable conditions the tendency to attain stable equilibrium can cause the molecules to rotate and set themselves in chains or sheets, with their maxima of attraction in line. The formation of sheets will commence at a great number of places throughout the solid, *i.e.*, wherever the initial random arrangement is sufficiently favourable. Evidently it is possible for the number of such "sufficiently favourable" arrangements to be enormously less than the total number of molecules, so that the ultimate result will be the formation of a number of units or groups, each containing a large number of molecules oriented according to some definite law. The relative arrangement of the units will, of course, be haphazard.

Now, in each unit there will, in general, be a direction which is, approximately at least, that of the minimum attractions of the majority of the molecules in the unit. Hence if the ratio of the maximum to the minimum attractions is sufficiently great, each unit can constitute a "flaw," and there appears to be no reason why the units should not be as large as the flaws have been shown to be in the case of glass. Thus, in order to explain the spontaneous weakening of glass, it is only necessary to suppose that the thermal agitation at about 1400° C. is sufficient to bring about the initial random formation.

It will be remembered that in the case of the freshly drawn fibres the reduction in tenacity required several hours for completion, so that the time taken was large compared with the time of cooling. Expressed in terms of molecular motion, this means that the molecules resist rotation very much more than they resist translation. This is in keeping with the conclusions of DEBYE,* who found that, on the basis of the quantum theory, the phenomena associated with the specific heat of solids could be explained only if the thermal vibrations of the molecules were regarded as practically irrotational. The same thing is shown more roughly, without introducing the quantum theory, by the law of DULONG and PETIT, which requires that each molecule shall have only three degrees of freedom.

The theory here put forward makes the spontaneous weakening a consequence of the attainment of a molecular configuration of stable equilibrium ; it therefore suggests that the weakening should be accompanied, in general, by a change in the dimensions of the solid. This has been verified by direct observation with a high-power microscope ; in the course of half an hour a spontaneously weakening glass fibre increased in length by about 0·1 per cent., while the length of a silica fibre decreased by about 0·03 per cent.

On account of the random arrangement of the molecular groups, this spontaneous change in unstrained volume must set up internal stresses, which may be sufficiently large to start cracks along the directions of least strength. In this connection it may be mentioned that irregularly shaped pieces of glass, of which some parts had been put into the strong unstable state by heating, have sometimes been observed to break spontaneously about an hour after cooling was practically complete.

It was remarked on p. 184 that cracks could not form spontaneously in a substance composed of molecules having spherical fields of force, as the process would involve an increase in potential energy. This is no longer true when the attraction is a function of orientation, as the surface energy of the cracks may be more than counterbalanced by the decrease in potential energy accompanying the molecular rearrangement.

For this reason, it is impossible to deduce the ratio of the maximum to the minimum molecular attractions from the ratio of the maximum and minimum strengths of the material, as it is possible that the spontaneous weakening is always accompanied by the formation of minute cracks, of the same size as the molecular groups.

It is probable that, in many cases, the most stable orientation of the molecules at a free surface is that in which their maxima of attraction lie along the surface. Such an orientation would in turn lead to a similar tendency on the part of the next layer of molecules, and so on, the tendency diminishing with increasing distance from the surface. There would therefore be a surface layer having the special property that in it the "flaws" ran parallel to the surface.

Hence this layer would be of exceptional strength in the direction of the surface. This suggests a reason for the experimental fact that the breaking load of wires and fibres consists mainly of two parts, one proportional to the area, and the other to the perimeter of the cross-section. The process of drawing, too, might predispose the molecules to take up positions with their maxima of attraction parallel to the surface.

---

* 'Ann. der Physik,' 39 (1912), p. 789.

If a perfectly clean glass plate be covered with gelatine and set aside, the gelatine gradually contracts, and as it does so it tears from the glass surface thin flakes up to about 0·06-inch diameter and shaped like oyster shells.* This tendency to flake at the surface is also observed when glass is broken by bending. This was particularly well shown in the specially prepared fibres used for the experiments described in the present paper. In almost all cases of flexural fracture the crack curled round on approaching the compression side, till it was nearly parallel to the surface. On two occasions the fracture divided before changing direction, the two branches going opposite ways along the fibre and a flake of length several times the diameter of the fibre was detached.

Surface flaking is also observed when some kinds of steel are subjected to repeated stress. Here the flakes are usually very small.

All these facts are evidently in complete agreement with the "surface layer" theory and, indeed, it is difficult to account for them on any other basis.

8. *Extended Application of the Molecular Orientation Theory.*

On the basis of the present theory, the physical properties of materials must be intimately related to the geometrical properties of the molecular sheet-formation. In order that a substance may exhibit the characteristic properties of crystals, it is clearly necessary for the sheets of molecules to be plane. In this case the crystals are, of course, the molecule groups or "units" referred to above. In "amorphous" materials, on the other hand, the sheets are probably curved.†

In materials of the former type, there must exist planes on which, if they are subjected to a sufficiently large shearing stress, the portions on either side of the planes can undergo a mutual sliding through a distance equal to any integral multiple of the molecular spacing, without fundamentally affecting the structure of the crystal. It is well known that the phenomenon of yield in crystals, and especially in metals, is of this nature. The planes in question are, of course, the well-known "gliding planes," and it is further possible that they may be identified also with the surfaces of least attraction. The stress at which gliding occurs in a single crystal must be determined in the following manner. The molecules of a crystal are normally in a configuration of stable equilibrium, and if two parts of the crystal slide on a gliding plane through one molecular space the resulting configuration is also stable. Between these two positions there must, in general, be one of higher potential energy, in which the equilibrium is unstable, and the shearing stress is determined by the condition that the rate at which work is done, in sliding from the stable to the unstable state, must be equal to the greatest rate of increase in potential energy which occurs during the passage between the two states. This rate will depend on the shape of the molecular fields of force, and may in particular cases be zero. Liquid crystals are doubtless of this type. The average shear stress, during yield, of a random aggregation of a large number of crystals, is doubtless greater than that of a single crystal, as the angle between the gliding planes and the maximum shear stress must vary from crystal to crystal and can be zero in only a few of them.

---

* Lord RAYLEIGH, 'Engineering,' 1917, vol. 103, p. 111, and H. E. HEAD, 'Engineering,' 1917, vol. 103, p. 138.

† See QUINCKE, 'Ann. der Physik,' (4), 46, 1915, p. 1025.

As the mutually gliding portions of a crystal pass from the stable to the unstable state, the molecular cohesion between them (normal to the gliding plane) must, in general, become less. In particular instances it may diminish to zero before the position of unstable equilibrium is reached. In these cases, shearing fracture along the gliding planes will occur, unless the material is subjected to a sufficiently high "hydrostatic" pressure, in addition to the shearing stress. Thus, a crystalline substance may be either ductile or brittle, according to nature of the applied stress. or it may be ductile at some temperatures and brittle at others, under the same kind of stress as has been actually observed by BENGOUGH and HANSON in the case of tensile tests of copper. This rupture in shear explains the characteristic fracture of short columns of brittle crystalline material under axial compression. The theory indicates that such fracture can always be prevented and yield set up by applying sufficient lateral pressure in addition to the longitudinal load ; this is in agreement with experiments on rocks such as marble and sandstone.* Conversely, a ductile substance might be made brittle if it were possible to apply to it a sufficiently large hydrostatic tension.

In the case of an alloy of, say, two metals A and B, suppose, as an example, that the sequence of molecules on either side of a gliding plane is

. A . B . B . A . B . B
B . B . A . B . B . A .

Let sliding occur (through one molecular space) to an adjoining position of stable equilibrium, or, say, to the configuration

. A . B . B . A . B . B . ←
. B . B . A . B . B . A →

Evidently, the structure in the neighbourhood of the gliding plane is in this case no longer the same as in the original crystal formation. It is therefore likely that the new state is one of higher potential energy, whence it is reasonable to suppose that the maximum rate of increase in potential, in sliding, is greater than it would have been had the potential of the two states been the same. Thus an alloy may be expected to have a higher yield-point than its most ductile constituent. This is in accordance with experience. For example, it is known that quenching from a high temperature hardens tool steel by preventing the separation of "ferrite," or iron containing no carbon.

---

* T. V. KARMAN, 'Zeitschr. Ver. Deutsch. Ing.,' 55, 1911.

In a single crystal the molecules are presumably in an equilibrium configuration of maximum stability. In this event, the equilibrium of molecules at or near intercrystalline boundaries, in a body composed of a large number of crystals, must, in general, be less stable than that of the molecules in the interior of the crystals. In fact, where the orientation of the component crystals is haphazard, the stability of the boundary molecules may be expected to range from the maximum of normal crystallisation down to zero, *i.e.*, neutral equilibrium. If such a body be subjected to a shear stress, some of the molecules in or near neutral equilibrium must, in general, become unstable, and these will tend to rotate to new positions of equilibrium. This rotation, however, will be strongly resisted, as has been seen, by forces doubtless of a viscous nature, and its amount will accordingly depend on the time during which the stress is applied. If, therefore, the strain is observed it will be found to increase slowly as time goes on, but at a constantly decreasing rate, as the molecules concerned approach equilibrium. If now the load is removed, these molecules must rotate in order to regain their original positions of equilibrium, and this process in turn will be retarded by viscous forces. Hence a small part of the observed strain will remain after the removal of the load, and this will gradually disappear as time goes on. These properties, known as "elastic after-working," are, of course, well known to belong to crystalline materials. Moreover, the theory shows that they should not be possessed by single crystals, and this has been demonstrated experimentally.*

There is a special type of gliding or yield which may occur at stresses below the normal yield point. Consider a pair of adjacent crystals, separated by a plane boundary. If these crystals are thought of as sliding relatively to each other, it will be seen that only in a finite number of the positions so taken up can the two be in stable equilibrium. Between each pair of such positions there must in general be one of unstable equilibrium. Suppose that, while near such an unstable position, the two crystals are embedded in a number of others. Under these conditions the boundary molecules of the two crystals will be pulled over in the direction of one or other of the two adjoining stable positions, and they will strain the solid in the process. If now the body is subjected to a shearing stress tending to cause relative displacement of the two crystals towards the other stable position, then at a certain value of this stress the molecules on either side of the boundary will be wrenched away, will pass through the position of instability, and will then take up a new position bearing the same relation to the second stable position as their original state did to the first. This new condition will, of course, persist after the removal of the load, as the original state cannot be regained without passing through unstable equilibrium, *i.e.*, a condition of maximum potential energy. To cause the crystals to pass through this condition it would be necessary to apply a load of opposite sign, and in this way the process might be repeated indefinitely. In a body composed of a large number of crystals there must be many arrangements of this type, in which adjacent crystals can execute inelastic oscillations about positions of unstable equilibrium, under alternating shear stresses below the ordinary yield stress. The consequent observed phenomena would correspond exactly with those known to be manifested in metals, under the name "elastic hysteresis."*

---

* H. v. WARTENBERG, 'Deutsch. Phys. Gesell., Verh.,' 20, pp. 113–122, August 30, 1918.

Experimentally, elastic hysteresis is distinguished from elastic after-working by the circumstance that it is completed very much more quickly. This is just what would be expected theoretically, on the view that molecular translation occurs much more readily than rotation.

It has been remarked that when a single crystal of a pure substance is caused to yield, its structure is fundamentally unaltered. This cannot hold, however, in the case of an aggregate of a large number of crystals arranged at random, or a crystal embedded in amorphous material. True, the material in the interior of each crystal can retain its original properties, but near the crystalline boundaries the structure must be violently distorted. As a result, it may be expected that the number of the molecules of inferior stability will be largely increased. Elastic after-working in metals should therefore be increased by overstraining or "cold-working." This, again, agrees with experience.

The foregoing considerations lend support to the view that each crystal of a severely cold-worked piece of metal is surrounded by an amorphous layer of appreciable thickness. If such a piece of metal undergoes a shear strain greater than that which can initiate yield in the normal crystalline substance, the average stress which is set up must be above the normal yield stress, for the part due to the amorphous layers must be the elastic stress corresponding with the strain, and this, by hypothesis, is greater than the yield stress. This part, moreover, will increase with the strain. It follows that yield in cold-worked metal should be less sharply defined, and should occur at a higher shear stress than in the normal crystalline variety. That this is actually the case is, of course, well known.

In the case of very large strains an important part of the shear stress must be taken by the amorphous boundary layers, and as a result the maximum tensile stress may reach a value sufficient to cause rupture of some favourably disposed crystals across their planes of least strength. This is, perhaps, the actual mode of rupture in ductile materials. On this view, the "ductility" of a metal depends simply on the relation between the tensile strength of the "flaws" and the normal yield stress. A substance whose ductility is small may still be "malleable," as hammering need not give rise to large tensile stresses.

The formation of non-crystalline material at the intercrystalline boundaries, when a piece of metal is over-strained, appears to provide an explanation of the sudden drop in stress which occurs immediately after the initiation of yield† in ductile metals.

---

* Guest and Lea, "Torsional Hysteresis of Mild Steel," 'Roy. Soc. Proc.,' A, June, 1917.

† Robertson and Cook, 'Roy. Soc. Proc.,' A, vol. 88, 1913, pp. 462–471.

Remembering that the surface tension of a substance is the work done in forming unit area of new surface, it will be seen that the tension of any surface of a crystal must depend on the angle it makes with the crystal axis. Thus the surface tension parallel to the planes of least strength must be less than that in any other direction. Consequently, in a body composed of a number of crystals there must exist a mutual surface tension at each intercrystalline boundary. Now, the theory of surface tension shows that the magnitude of such a mutual tension is greatly diminished by making the transition between the two bodies more gradual. Hence the formation of the amorphous boundary layer involves a reduction in the surface energy of the crystals, and this is shown in the experiments by a drop in the stress. If this account of the phenomenon is complete, the drop in stress must be determined by the condition that the loss of strain energy equals the reduction in surface energy. The mechanism of the process appears to be that the breaking up of the boundary, which must accompany yield, is resisted by the surface tension, and yielding therefore requires a higher stress for its initiation than for its maintenance.

According to this view, the loss of strain energy should be inversely proportional to the linear dimensions of the crystals. Hence the results of different experiments should show considerable variation in the magnitude of the drop in stress. This is actually the case ; a single series of experiments on mild steel, by ROBERTSON and COOK, gave drops varying from 17 per cent. to 36 per cent., while in other experiments as little as 7 per cent. has been observed.

In the above series of experiments the average loss of strain energy was about 12 inch-lbs. per cubic inch. Assuming, for simplicity, that the crystals were cubes, of, say, 0·001-inch side (which is a fair value for well-treated mild steel), the area of the intercrystal surface was 3000 sq. inches per cubic inch. These figures give the average intercrystal surface tension as 0·004 lbs. per inch. This is certainly of the right order of magnitude.

Many of the phenomena discussed above will be more complicated, in practice, if the coefficient of expansion of the crystals is not the same in all directions. In such an event, internal stresses will be set up in cooling, on account of the random arrangement of the crystals, and these stresses must be taken into consideration in applying the theory.

There remains for consideration the problem of the fracture of metals under alternating stress. It is known that fatigue failure occurs as the result of cracking after repeated slipping on gliding planes, and the theory has been advanced* that this cracking is due to repeated to and fro sliding and consequent attrition and removal of material from the gliding planes. This theory presents some difficulties, in that it does not explain how the attrition can occur, or the method of disposing of the debris.

A theory which is free from these objections may be constructed if it is supposed that a change in volume occurs on the passage of the metal from the crystalline to the amorphous state. This assumption is, of course, known to be valid for many substances

* EWING and HUMFREY, 'Phil. Trans.,' A, 1902, p. 200.

at their melting points, but at lower temperatures there seems to be no definite information available.

This assumption being granted, suppose that a piece of material which contracts on decrystallising is being subjected to a stress cycle just sufficient to cause repeated slipping in the most favourably disposed crystals. As a result, the material at the boundaries of these crystals will become amorphous, and the quantity of amorphous material will increase continuously as long as the repeated slipping goes on. But, by hypothesis, the unstrained volume of the amorphous phase is less than the space it filled when in the crystalline state. Hence all the material in the immediate neighbourhood will be subjected to a tensile stress, and as soon as this exceeds a certain critical value a crack will form. It has been observed above that the application of a sufficiently large hydrostatic tension may be expected to make a ductile substance brittle. Hence the crack may occur either in tension or in shear, according to the properties of the material and the nature of the applied stress. Further alternations of stress will cause this crack to spread until complete rupture occurs. This theory makes the limiting safe range of stress equal to that which just fails to maintain repeated sliding in the most favourably disposed crystals.

It may be asked why such cracking does not take place in a static test where the quantity of amorphous material, once yield has fairly started, is presumably much greater. The answer to this is two-fold. In the first place, if the material becomes amorphous round all, or nearly all, the crystals of a piece of metal, it is evident that it will contract as a whole and no great tensile stress will be set up. In the case where only a few crystals yield, the tension arises from the rigidity of the unchanged surrounding metal.

In the second place, even if some crystals do crack, the cracks will not, in general, tend to spread through the ductile cores of the neighbouring crystals, unless the applied load is alternating, on account of the equalisation of stress due to yield.

The safe limit of alternating stress will usually be less than the apparent stress necessary to initiate yield in a static test, on account of initial stresses, including those due to unequal contraction of the crystals.

The theory indicates that the cracking of the first crystal marks a critical point in the history of the piece. At any earlier stage the effects of the previous loading may be removed by heat treatment, or possibly by a rest interval, but once a crack has formed this cannot be done. True, the tension may be relieved and the ruptured crystal may even be compressed somewhat, but this cannot, in general, close the crack, as cracking is not a "reversible" operation. An exception may occur if the top temperature of the heat-treatment is sufficient to bring the molecules on either side of the crack within mutual range by thermal agitation, but it is unlikely that this can happen save in the case of very small cracks.

If this theory is correct, it appears at first sight that the phenomenon of fatigue failure must be confined to substances which contract on decrystallising. This, however is not necessarily so. If, for instance, a small thickness of material at the interface between two crystals were to increase in volume, it could not be said without proof that tensile stresses would not be set up thereby, in addition to compressions. In some cases, in fact, it is obvious that there must be tensions. Thus, if the outer layer of a sphere increases in volume, the matter inside must be subjected to a tensile stress.

The effect of overstrain on the density of metals is at present under investigation at the Royal Aircraft Establishment. The work is not yet sufficiently complete for detailed publication, but it may be mentioned here that the expected change in density has been found, and that the results already obtained are such as to leave little room for doubt that this change is in fact the cause of fatigue failure in metals. Thus, in overstraining mild steel by means of a pure shearing stress, a decrease in average density of as much as 0·25 per cent. has been observed.

Some progress has also been made in the direction of estimating the internal stresses set up as a result of the change in density, and it has been found that an average change of the magnitude mentioned above could give rise to a hydrostatic tensile stress in the cores of the crystals, of the order of 30,000 lbs. per sq. inch.

Dealing now with materials whose molecular sheet-formations are curved, it is at once evident that all yield, or slide, phenomena must be absent, as possible gliding planes do not exist. Thus, this case, though geometrically more complicated, is practically much simpler than that in which the sheets are plane. The theoretical properties of materials having the curved type of formation appear to correspond exactly with those known to belong to brittle "amorphous" substances. Exactly as in the case of crystalline materials, elastic after-working is explained by the inferor stability of molecules near the boundaries of the units of molecular configuration, but elastic hysteresis should not occur. If adequate precautions are taken to avoid secondary tensile stresses, fracture of short columns in compression should occur at stresses of an altogether higher order than in the case of crystals. In this connection it may be remarked that the compressive strength of fused silica is about 25 times as great as its ordinary tensile strength.

It appears from the foregoing discussion that the molecular orientation theory is capable of giving a satisfactory general account of many phenomena relating to the mechanical properties of solids, though closer investigation will perhaps show that the agreement is in some cases superficial only. Such questions as the effects of unequal cooling, foreign inclusions and local impurities, and the behaviour of mixtures of different crystals, have not been dealt with; it is thought that these are matters of detail whose discussion cannot usefully precede the establishment of the general principles on which they depend.

9. *Practical Limitations of the Elastic Theory.*

It is now possible to indicate the directions in which the ordinary mathematical theory of elasticity may be expected to fail when applied to real solids.

It is a fundamental assumption of the mathematical theory that it is legitimate to replace summation of the molecular forces by integration. In general this can only be true if the smallest material dimension, involved in the calculations, is large compared with the unit of structure of the substance. In crystalline metals the crystals appear, from the foregoing investigation, to be anisotropic and they must therefore be regarded as the units of structure. Hence the theory of isotropic homogeneous solids may break down if applied to metals in cases where the smallest linear dimension involved is not many times the length of a crystal.

Similar considerations apply to solids such as glass, save that here the units of structure are probably curved.

The most important practical case of failure is that of a re-entrant angle or groove. Here the theory may break down if the radius of curvature of the re-entering corner is but a small number of crystals long. An extreme instance is that of a surface scratch, where the radius of curvature may be but a fraction of the length of the crystals.

In the case of brittle materials the general nature of the effect of scratches on strength may be inferred from the theoretical criterion of rupture enunciated in section 2 above. Whether the material be isotropic or anisotropic, homogeneous or heterogeneous, it is necessary on dimensional grounds that the strain energy shall depend on a higher power of the depth of the scratch than the surface energy. It follows that small scratches must reduce the strength less than large ones of the same shape. Hence, where the tenacity of the material, under "uniform" stress, is determined by the presence of "flaws," it must always be possible to find a certain depth of scratch whose breaking stress is equal to that of the flaws. Evidently such a scratch can have no influence on the strength of the piece. Deeper scratches must have some weakening effect, which must increase with the depth, until in the limit the strength of very large grooves may be found by means of the elastic theory and the appropriate empirical hypothesis of rupture.

In the case of ductile metals, the effect of scratches is important only under alternating or repeated stresses. On the theory advanced in the preceding section, fatigue failure under such stresses is determined by phenomena which occur at the intercrystalline boundaries. Hence the strength of a scratched piece is fixed, not by the maximum stress range in the corner of the scratch, but by the stress range at a certain distance below the surface. This distance cannot be less than the width of one crystal, and it may be greater. Elastic theory suggests that the stress due to a scratch falls off very rapidly with increasing distance from the re-entrant corner, so that the relatively small effect of scratches in fatigue tests is readily explained.

Possibly many published results bearing on this matter depend more on initial skin stresses than on sharp corner effects.

10. *Methods of Increasing the Strength of Materials.*

The most obvious means of making the theoretical molecular tenacity available for technical purposes is to break up the molecular sheet-formation and so eliminate the "flaws." In the case of crystalline material this has the further advantage of eliminating yield and probably also fatigue failure.

In materials which normally have curved sheets, the molecular fields of force are presumably asymmetrical, and the process indicated above would lead of necessity to a random arrangement, which might be unstable. It has been seen that in glass and fused silica it is actually unstable, except in the case of the finest fibres.

As regards crystalline materials, however, in which the fields of force must have some sort of symmetry, there seems to be no reason why there should not be possible a very fine grained stable configuration, which could be derived from the ordinary crystalline form by appropriate rotations of certain molecules to new positions of stable equilibrium, in such a way as to break up the gliding planes. The grain of such a structure need be but a few molecules long, and its strength would approximate to the theoretical value corresponding with the heat of vaporisation.

There is some evidence that mild steel which has been put into the amorphous condition by over-strain tends, under certain conditions, to take up a stable fine-grained formation of this kind, in preference to resuming its original coarse crystalline configuration, in that a temperature of 0° C. appears to prevent recovery from tensile over-strain.*

These considerations suggest that if a piece of metal were rendered completely amorphous by cold-working, and then suitably heat-treated, its molecules might take up the stable strong configuration already described. The theory indicates, however, that over-straining tends to set up tensile stresses in the unchanged parts of the crystals which may start cracks long before decrystallisation is complete. Such cracking could be prevented if the over-straining were carried out under a sufficiently great hydrostatic pressure, and this line of research seems to be well worth following up. It might, of course, be found that the requisite pressure was so enormous as to render the method unworkable, but if the theory is sound there seems to be no other reason why definite results should not be obtained.

The problem may be attacked in another way. As has been seen, the theory suggests that the drop in stress at the initiation of yield is due to the surface energy of the inter-crystal boundaries. Thus the yield point may be raised by "refining" the metal, *i.e.*, so heat-treating it as to reduce the size of the crystals. The limit of refinement is, doubtless, reached when each "crystal" contains but a single molecule and the material is then in the strong stable state already described.

Refining is also of great value in connection with resistance to fatigue failure. Suppose, in accordance with the foregoing theory of fatigue, that one crystal has been fractured, then the general criterion of rupture shows that the crack cannot spread unless the material is subjected to a certain minimum stress, which is greater the smaller the crack. Thus, reducing the size of the crystals increases the stress necessary to cause the initial crack to extend. There is therefore a critical size of crystal for which the stress-range necessary to spread the crack is equal to that necessary to start it. Until the refining has reached that stage it can have no effect on the magnitude of the safe stress range, but from that point on the range must increase progressively with refinement until the limit is reached, as before, when each "crystal" contains but one molecule.

---

* COKER, 'Phys. Rev.,' 15, August, 1902.

It therefore appears that refining is one avenue of approach towards the ideal state of maximum strength. Strangely enough, another line of argument suggests that the reverse of refinement might be effective in securing the desired result, in certain special cases. If a wire is required to withstand a simple tension, it seems that the best arrangement is that in which the strongest directions of all the molecules are parallel to the axis of the wire. This is equivalent to making the wire out of a single crystal. The theoretical tenacity would not be obtained, however, if the gliding planes made with the axis angles other than 0° C. and 90° C., as yield would occur.

If, in passing from the normal crystalline to the strong fine-grained state, the necessary orientation of the molecules were performed in accordance with some regular plan, the resulting configuration would possess some kind of symmetry, and the material might therefore exhibit crystalline properties. In cases where a substance exists in nature in several different crystalline forms, of which one is much stronger than the others, it may be that the strong modification is of the fine-grained type here considered. Thus, diamond may be a fine-grained modification of carbon. If this view is correct, it suggests that the transformation of carbon into diamond requires, firstly, the existence of conditions of temperature and pressure under which diamond has less potential energy than carbon; and, secondly, the provision of means for causing relative rotation of the molecules. In the attempts which have so far been made in this direction, attention seems to have been concentrated on satisfying the former requirement, the possible existence of the latter one having been overlooked. The most obvious way of satisfying it, if the mechanical difficulties could be overcome, would appear to be the application of suitable shearing stresses in addition to the hydrostatic pressure.

11. *Application of the Theory to Liquids.*

A detailed discussion of the properties of liquids, in the light of the present theory, would scarcely fall within the scope of this paper. One prediction which has been made, however, and which has been verified experimentally, affords such a remarkable confirmation of the general theory that it is felt that no apology is necessary for introducing it here. Consider a solid composed of molecules whose attraction is a function of orientation, the molecules being arranged in groups, in accordance with the theory outlined in the preceding pages. If the temperature of this body be supposed to be increasing, it will be seen that at some temperature the kinetic reactions due to the thermal vibrations must overcome the minimum attractions of the molecules in each group. It is clear, therefore, that at this temperature the substance will be unable to withstand shearing stresses. At the same time it cannot vaporise, as the molecules must still be held together in chains by their maximum attractions. In other words, the transformation which has been discussed is simply the liquefaction of the solid.

This view of the phenomenon of melting indicates that the molecules of liquids are in general arranged in groups or chains, of a length comparable with that of the structure ascribed to solids in the preceding work, or, say, 10^4 molecules.

If, therefore, a liquid be contained in a solid boundary which it wets, the ends of these chains may be expected to attach themselves to the solid; and if at any point the distance between the bounding walls is less than the length of the chains, some of the latter will attach themselves to both walls and hinder the free flow of the liquid and the relative movement, if any, of the boundaries. At such a point the liquid will act as a solid under any stress which is insufficient to break the chains.

This has been verified experimentally. The apparatus consisted of a polished steel ball 1 inch in diameter, and a block of hard tool steel containing a circular hole about 4 inches long. The hole was carefully ground, after hardening, to a diameter about 0·0001 inch greater, at its smallest part, than the diameter of the ball. When both were dry the ball passed freely through the hole. If, however, they were wetted with a liquid, considerable pressure was necessary to force the ball through. This resistance possessed the characteristic "stickiness" of solid friction, and was exactly the kind of resistance which would have been expected in forcing the dry ball through a hole which was too small for it.

To show that the resistance was a true "solid stress" and not due merely to viscosity, the apparatus was on one occasion left for a week, with the weight of the ball supported by the stress in the liquid (paraffin oil). The hole was vertical, so that there was no normal pressure between its surface and the surface of the ball. During this period no motion whatever could be detected.

It is essential to the success of these experiments that the ball and hole should be thoroughly wetted by the liquid. For this reason the liquids used have been chiefly paraffin oil and lubricating oils, but on one occasion the effect was obtained with water.

The present theory suggests a reason for the very low tensile strength of liquids. If a liquid is composed of a random aggregation of chains of molecules, it may reasonably be expected to contain regions of dimensions comparable with, but smaller than, the length of the chains, across which no chains run. Rupture of the liquid will evidently occur by the enlargement of these cavities. Now the tension, R, necessary to enlarge a spherical cavity of diameter, D, in a liquid of surface tension, T, is given by

$$R = 4T/D.$$

In the case of water, the tensile strength, R, is about 70 lbs. per sq. inch at ordinary temperatures, while T is about 0·00042 lbs. per inch. Hence the cavities, if spherical, must be at least 0·000024 inch in diameter. This is of the order indicated by the theory.

The foregoing conclusions are of especial interest in their relation to the theory of ROSENHAIN,* on which many of the properties of metals, and particularly "season cracking" under prolonged stress, are explained by supposing that the crystals are cemented together by very thin layers of amorphous material having the properties of an extremely viscous undercooled liquid. The experiments described above show that fluidity is not a property which can be ascribed *a priori* to such films. Hence if the view of ROSENHAIN and ARCHBUTT were to be definitely established, it would be necessary to regard it, not as a theory of season cracking in terms of the known properties of materials, but as a deduction of the properties of the intercrystalline layers from the phenomena of season cracking. Looked at in this way, it would be of extreme interest, for it would show that the molecular arrangement of the intercrystalline layers could not be of the coarse-grained type characteristic of the normal states of solids and liquids.

It is clear that the foregoing theory of liquids is not free from objection, and that in some respects it appears to be less satisfactory than existing theories. The most obvious objection is that it seems to be incompatible with accepted determination of the molecular weight of liquids. Since, however, these experiments are based ultimately on kinetic considerations, the author believes that this difficulty will not in fact arise unless the requisite bonds between the molecules of each group are found to be sufficiently strong to cause appreciable modification of the average molecular kinetic energy.

12. *Summary of Conclusions.*

(1) The ordinary hypothesis of rupture cannot be employed to predict the safe range of alternating stress which can be applied to metal having a scratched surface. The safe range of an unscratched test piece appears to be slightly less than the yield range, but if the surface is scratched the safe range may be several times the range which causes yield in the corners of the scratches.

(2) The "theorem of minimum potential energy" may be extended so as to be capable of predicting the breaking loads of elastic solids, if account is taken of the increase of surface energy which occurs during the formation of cracks.

(3) The breaking load of a thin plate of glass having in it a sufficiently long straight crack normal to the applied stress, is inversely proportional to the square root of the length of the crack. The maximum tensile stress in the corners of the crack is more than ten times as great as the tensile strength of the material, as measured in an ordinary test.

(4) The foregoing observation is in agreement with the known fact that the observed strength of materials is less than one-tenth of the strength deduced indirectly from physical data, on the assumption that the materials are isotropic. The observed

---

* W. ROSENHAIN and D. EWEN, "Intercrystalline Cohesion in Metals," 'J. Inst. Metals,' vol. 8 (1912); and W. ROSENHAIN and S. L. ARCHBUTT, "On the Intercrystalline Fracture of Metals under Prolonged Application of Stress (Preliminary Paper)," 'Roy. Soc. Proc.,' A, vol. 96 (1919).

strength is, in fact, no greater than it would be, according to the theory, if the test pieces contained cracks several thousand molecules long.

(5) It has been found possible to prepare rods and fibres of glass and fused quartz which have a tenacity of about one million pounds per square inch (approximately the theoretical strength) when tested in the ordinary way. The strength so observed diminishes spontaneously, however, to a lower steady value, which it reaches a few hours after the fibre has been prepared. This steady value depends on the diameter of the fibre. In the case of large rods it is the same as the ordinary tenacity, whereas in the finest fibres the strength diminishes but little from its initial high value. The relation between diameter and strength is of practically the same form for glass fibres as for metal wires.

(6) If it is assumed that intermolecular attraction is a function of the relative orientation of the attracting molecules, it is possible to construct a theory of all the phenomena mentioned in (3), (4) and (5) above. In the case of crystalline substances the theory also appears to explain yield and shearing fracture; elastic hysteresis; elastic afterworking; the fracture in tension of ductile materials and the flow of brittle materials under combined shearing stress and hydrostatic pressure; the drop in stress which occurs on the initiation of yield in ductile substances; fatigue failure under alternating stress; and the relatively slight effect of surface scratches on fatigue strength. In the case of non-crystalline materials the theory explains elastic afterworking and the great strength of short columns in compression.

(7) The theory shows that the application of the mathematical theory of homegeneous elastic solids to real substances may lead to error, unless the smallest material dimension involved, *e.g.*, the radius of curvature at the corner in the case of a scratch, is not many times the length of a crystal.

(8) It should be possible to raise the yield point of a crystalline substance by "refining" it, until at the ultimate limit of refinement the yield stress should be of the same order as the theoretical strength. It should also be possible similarly to increase the tenacity. Up to a certain stage the fatigue range should be unaffected by refining, but thereafter it should increase in the same degree as resistance to static stress.

(9) The theory requires that a thin film of liquid enclosed between solid boundaries which it wets should act as a solid. Experimental confirmation of this has been obtained.

In conclusion, the author desires to place on record his indebtedness to many past and present members of the staff of the Royal Aircraft Establishment for their valuable criticism and assistance, and also to Prof. C. F. JENKIN, at whose request the work on scratches was commenced.

[*Note.*—It has been found that the method of calculating the strain energy of a cracked plate, which is used in Section 3 of this paper, requires correction. The correction affects the numerical values of all quantities calculated from equations (6), (7), (8), (10), (11), (12) and (13), but not their order of magnitude. The main argument of the paper is therefore not impaired, since it deals only with the order of magnitude of the results involved, but some reconsideration of the experimental verification of the theory is necessary.]

VII. *Reduction of Error by Linear Compounding.*

By W. F. SHEPPARD, *Sc.D., LL.M.*

Communicated by E. T. WHITTAKER, *F.R.S.*

Received April 23,—Read June 24, 1920.

CONTENTS.

Page
1. Introductory 200

CONJUGATE SETS.

2. Conjugate set 201
3. Relations between original set and conjugate set 203
4. Two related pairs of conjugate sets 204
5. Sums as conjugates of differences 205

REDUCTION OF ERROR (GENERAL).

6. General theorems 208
7. Notation: and particular values. 210
8. Improved values in terms of conjugates 211
9. Mean products of error of improved values 212
10. Expression in terms of a related set 214
11. Special case of differences 215

RELATION OF "REDUCTION OF ERROR" TO "FITTING" (OF A POLYNOMIAL).

12. Standard system 215
13. Fitting by least squares 216
14. Fitting by moments 219
15. Reduction of error 219
16. Difference of the two processes 219

SOME STEPS IN THE GENERAL SOLUTION.

17. Preliminary 220
18. Formula of progression 220
19. Mean products of error (alternative formula) 221

APPLICATION TO SELF-CONJUGATE SET.

20. Preliminary 222
21. Improved values 225
22. Extent of improvement (central differences) 230
23. Smoothing 231

[Published October 22, 1920.

APPENDICES.

Page
APPENDIX I.—The correlation-determinant 232
„ II.—Frequency of correlated errors 233
„ III.—Improved advancing differences in terms of sums 235
„ IV.—Formulæ in terms of u's. 236

1. *Introductory.*—This paper is a development of two earlier papers,[*] which for brevity I call "Reduction" and "Fitting" respectively. The paper[†] immediately preceding "Fitting" is referred to as "Factorial Moments."

These earlier papers deal with two problems, which are closely connected and have the same solution. For both of them, the data are a set of quantities u_0, u_1, u_2, ... of the same kind, which we regard as representing certain true values U_0, U_1, U_2, ..., with errors e_0, e_1, e_2, ..., so that $u_r = U_r + e_r$. These errors may be independent or may be correlated in any way. The first problem is based on the assumption (which defines the class of cases we are dealing with) that the sequence of U's is fairly regular, so that differences after those of a certain order, which we will call j, are negligible. This being so, we may alter any u, or any linear compound of the u's, such as an interpolation-formula, by adding to it any linear compound of the negligible differences. (I use the term "linear compound" in preference to "linear function," since there is no consideration of functionality.) The problem is to find the value of the resulting sum when, by suitable choice of the coefficients in the added portion, the mean square of error of the sum is a minimum. This is the problem of "reduction of error." For the second problem it is assumed that U_r is a polynomial in r of degree j, and the problem is to find the coefficients in this polynomial by the method of least squares. This is the problem of "fitting."

The practical solution of these problems for the general case, in which the errors are correlated, is not easy. The particular case which is simple is that in which the errors all have the same mean square, which by a suitable choice of unit is taken to be 1, and the mean products of error are all 0. (In the previous papers I have called this system of errors the *standard* system; in the present paper the set of u's which possesses this property is called a *self-conjugate* set.) In "Reduction" I gave the solution for this particular case in terms of central differences, and in "Fitting" I gave the solution in terms of advancing differences and of advancing and central sums, formed in a particular way. I also gave expressions in terms of the u's, but these were rather complicated. It remained to obtain expressions for the mean squares of error of the new values, in order to compare them with those of the old values. In doing this I found that the whole of the work could be very much simplified by using certain general theorems, which applied not only to the special case of the standard system but also to the general case, and even to a still more general problem in which, in the one aspect, the reduction of error is effected by means of quantities which are not necessarily a set of differences, or in which, in the other aspect, U_r is not necessarily a polynomial in r of degree j but is a linear compound, with coefficients to be determined, of any $j+1$ functions of r; and the present paper is mainly concerned with these general theorems, so that to a certain extent it supersedes the previous papers.

---

[*] "Reduction of Errors by means of Negligible Differences," 'Fifth International Congress of Mathematicians,' Cambridge, 1912, ii., 348–384; "Fitting of Polynomial by Method of Least Squares," 'Proceedings of the London Mathematical Society,' 2nd series, xiii., 97–108.

[†] "Factorial Moments in terms of Sums or Differences," 'Proceedings of the London Mathematical Society,' 2nd series, xiii., 81–96.

The abbreviations l.c., m.s.e., m.p.e., are used for linear compound, mean square of error, mean product of errors. The mean square of error of A is denoted by $(A\ ;A)$, and the mean product of errors of A and B by $(A\ ;B)$ or $(B\ ;A)$. Other special notations used in the paper are the same as in the three papers mentioned at the beginning of this section, or are explained in §§ 3, 5 (iii.), 7, 17, and 20.

CONJUGATE SETS.

2. *Conjugate Set.*—(i.) Let A, B, C, D, ... be a set of quantities, not necessarily all of the same kind, containing coexistent errors which are either independent or correlated in any way. For the purpose of the following investigations it is convenient to consider, in connexion with these quantities, another set of quantities, G, H, J, K, ... , equal to them in number and connected with them by the conditions that (1) each quantity of the second set is a l.c. of those of the first set, and (2) the m.p.e. of corresponding members of the two sets is 1 and that of members which do not correspond is 0. If we replace the quantities of the two sets by A_0, A_1, A_2, ... , and G_0, G_1, G_2, ... , we can express this latter condition by saying that m.p.e. of G_r and $A_s = 0\ (s \neq r)$ or $1\ (s = r)$. The second set of quantities is said to be *conjugate* to the first.

(ii.) Let the member of the second set which corresponds to C of the first set be J. To determine J, let us write

$$J = aA+bB+cC+dD+\ldots.$$

Then, denoting the m.p.e. of A and B by $(A\ ;B)$, condition (2) gives

$$\begin{aligned}
(A\ ;A)\,a+(A\ ;B)\,b+(A\ ;C)\,c+(A\ ;D)\,d+\ldots &= 0,\\
(B\ ;A)\,a+(B\ ;B)\,b+(B\ ;C)\,c+(B\ ;D)\,d+\ldots &= 0,\\
(C\ ;A)\,a+(C\ ;B)\,b+(C\ ;C)\,c+(C\ ;D)\,d+\ldots &= 1,\\
(D\ ;A)\,a+(D\ ;B)\,b+(D\ ;C)\,c+(D\ ;D)\,d+\ldots &= 0,
\end{aligned}$$

&c.

There are as many equations as there are coefficients a, b, c, d, ... ; and the values of these are thus uniquely determined.

(iii.) The values of a, b, c, ... as found from the above equations have as their denominator the determinant

$$\Theta \equiv \begin{vmatrix} (A\,;A) & (A\,;B) & (A\,;C) & (A\,;D) & \ldots \\ (B\,;A) & (B\,;B) & (B\,;C) & (B\,;D) & \ldots \\ (C\,;A) & (C\,;B) & (C\,;C) & (C\,;D) & \ldots \\ (D\,;A) & (D\,;B) & (D\,;C) & (D\,;D) & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{vmatrix}.$$

There is therefore no conjugate set if this determinant is zero. The nature of the relations which in this case hold between the errors is considered in Appendix I., § 3.

(iv.) Since the members of the conjugate set are l.cc. of those of the original set, the converse also holds. Regrouping the equations which determine the coefficients, it will be seen that the original set is conjugate to the conjugate set; *i.e.*, that the two sets are conjugate to each other. The formulæ for the members of the original set in terms of those of the conjugate set are

$$\left.\begin{aligned} A &= (A\,;A)\,G+(A\,;B)\,H+(A\,;C)\,J+\ldots, \\ B &= (B\,;A)\,G+(B\,;B)\,H+(B\,;C)\,J+\ldots, \\ C &= (C\,;A)\,G+(C\,;B)\,H+(C\,;C)\,J+\ldots, \\ &\text{\&c.} \end{aligned}\right\} \quad \ldots \ldots \quad (1)$$

These follow from the solution of the equations in (ii.), by the ordinary properties of determinants; or they may be obtained more simply by determining the coefficients of G, H, J, ... in each case from the second of the conditions stated in (i.).

(v). By means of these relations we can not only express any l.c. of the quantities of either set in terms of those of the other set, but we can also express any such l.c. in terms of particular quantities of one set and those of the conjugate set which correspond to the remaining quantities. We can, for instance, express any l.c. of G, H, J, K, ... in terms of A, B, J, K, ... by using the first two equations in (1) to determine G and H in terms of A, B, J, K, ... The results involve a certain determinant in the denominator; it is shown in Appendix I., § 4, that this is not zero if Θ is not zero.

(vi.) Two special cases may be mentioned :—

(a) If the errors of A, B, C, D, ... form a standard system, *i.e.*, if the m.s.e. of each of the quantities is 1 and the m.p.e. of each pair of quantities is 0, the conjugate set is identical with the original set; and conversely. A set which is identical with the conjugate set will be called a *self-conjugate* set.

(*b*) If the m.p.e. of each pair of quantities of the original set is 0, but the m.ss.e. are not all 1, this is also the case for the conjugate set. The original set being $A, B, C, \ldots$, the quantities of the conjugate set are $A/(A\,;A)$, $B/(B\,;B)$, $C/(C\,;C)$, ... ; and their m.ss.e. are $1/(A\,;A)$, $1/(B\,;B)$, $1/(C\,;C)$, ...

3. *Relations between Original Set and Conjugate Set.*—For expressing a member of either set in terms of the members of the other set, it is convenient to give them a linear order. We therefore denote the members of the original set by $\delta_0, \delta_1, \delta_2, \ldots \delta_l$ and those of the conjugate set by $\sigma_0, \sigma_1, \sigma_2, \ldots \sigma_l$. Also we write

$$\zeta_{r,t} \equiv \text{m.p.e. of } \delta_r \text{ and } \delta_t = \zeta_{t,r}, \qquad (2)$$

$$\eta_{r,t} \equiv \text{m.p.e. of } \sigma_r \text{ and } \sigma_t = \eta_{t,r}. \qquad (3)$$

(i.) The condition of conjugacy is that ($r = 0, 1, 2, \ldots l$; $t = 0, 1, 2, \ldots l$)

$$\text{m.p.e. of } \delta_r \text{ and } \sigma_t = 0\,(r \neq t) \quad\text{or}\quad 1\,(r = t). \qquad (4)$$

(ii.) The expression for δ_r in terms of the σ's is (*cf.* § 2 (iv.))

$$\delta_r = \zeta_{r,0}\sigma_0 + \zeta_{r,1}\sigma_1 + \zeta_{r,2}\sigma_2 + \ldots + \zeta_{r,l}\sigma_l. \qquad (5)$$

[For, if we write

$$\delta_r = a_0\sigma_0 + a_1\sigma_1 + a_2\sigma_2 + \ldots + a_l\sigma_l,$$

then (2) and (4) give

$$\begin{aligned}\zeta_{r,t} &= \text{m.p.e. of } \delta_t \text{ and } a_0\sigma_0 + a_1\sigma_1 + \ldots + a_l\sigma_l \\ &= a_t.]\end{aligned}$$

(iii.) Similarly the expression for σ_t in terms of the δ's is

$$\sigma_t = \eta_{0,t}\delta_0 + \eta_{1,t}\delta_1 + \eta_{2,t}\delta_2 + \ldots + \eta_{l,t}\delta_l. \qquad (6)$$

(iv.) The relations between the ζ's and the η's are easily deduced from (5) and (6). If we write

$$Z \equiv \begin{vmatrix} \zeta_{0,0} & \zeta_{0,1} & \zeta_{0,2} & \cdots & \zeta_{0,l} \\ \zeta_{1,0} & \zeta_{1,1} & \zeta_{1,2} & \cdots & \zeta_{1,l} \\ \zeta_{2,0} & \zeta_{2,1} & \zeta_{2,2} & \cdots & \zeta_{2,l} \\ \vdots & \vdots & \vdots & & \vdots \\ \zeta_{l,0} & \zeta_{l,1} & \zeta_{l,2} & \cdots & \zeta_{l,l} \end{vmatrix}, \qquad (7)$$

$$Z_{p,q} \equiv \text{cofactor of } \zeta_{p,q} \text{ in } Z = Z_{q,p}, \qquad (8)$$

$$H \equiv \begin{vmatrix} \eta_{0,0} & \eta_{0,1} & \eta_{0,2} & \dots & \eta_{0,l} \\ \eta_{1,0} & \eta_{1,1} & \eta_{1,2} & \dots & \eta_{1,l} \\ \eta_{2,0} & \eta_{2,1} & \eta_{2,2} & \dots & \eta_{2,l} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \eta_{l,0} & \eta_{l,1} & \eta_{l,2} & \dots & \eta_{l,l} \end{vmatrix}, \qquad (9)$$

$$H_{p,q} \equiv \text{cofactor of } \eta_{p,q} \text{ in } H = H_{q,p}, \qquad (10)$$

then

$$\eta_{p,q} = Z_{p,q}/Z, \qquad (11)$$

$$\zeta_{p,q} = H_{p,q}/H, \qquad (12)$$

$$HZ = 1. \qquad (13)$$

(v.) The assumption that there is a conjugate set implies (*cf.* § 2 (iii.)) that Z is not = 0. It follows from (13) that H is not = 0. It also follows (see Appendix I., § 4 (*b*)) that none of the principal minors of Z or of H are = 0.

4. *Two Related Pairs of Conjugate Sets.*—(i.) Suppose that there is another set of $l+1$ quantities $u_0, u_1, u_2, \dots u_l$, connected with the δ's by the linear relations $(r = 0, 1, 2, \dots l)$

$$u_r = (r_0)\,\delta_0 + (r_1)\,\delta_1 + (r_2)\,\delta_2 + \dots + (r_l)\,\delta_l. \qquad (14)$$

Then, by the condition of conjugacy of the δ's and the σ's,

$$(r_t) = \text{m.p.e. of } u_r \text{ and } \sigma_t. \qquad (15)$$

Let the set conjugate to $u_0, u_1, u_2, \dots u_l$ be $y_0, y_1, y_2, \dots y_l$. Then there are linear relations between the y's and the u's and between the σ's and the δ's, and therefore also, by (14), between the y's and the σ's. To find the σ's in terms of the y's, we write (15) in the form

$$(r_t) = \text{m.p.e. of } \sigma_t \text{ and } u_r;$$

and we see that $(t = 0, 1, 2, \dots l)$

$$\sigma_t = (0_t)\,y_0 + (1_t)\,y_1 + (2_t)\,y_2 + \dots + (l_t)\,y_l. \qquad (16)$$

(ii.) Similarly, if the expression for the δ's in terms of the u's is $(s = 0, 1, 2, \dots l)$

$$\delta_s = \{s_0\}\,u_0 + \{s_1\}\,u_1 + \{s_2\}\,u_2 + \dots + \{s_l\}\,u_l, \qquad (17)$$

where, by (14),

$$(r_0)\,\{0_t\} + (r_1)\,\{1_t\} + \dots + (r_l)\,\{l_t\} = 0\;(r \neq t) \quad \text{or} \quad 1\;(r = t), \qquad (18)$$

$$\{s_0\}\,(0_t) + \{s_1\}\,(1_t) + \dots + \{s_l\}\,(l_t) = 0\;(r \neq t) \quad \text{or} \quad 1\;(r = t), \qquad (19)$$

then

$$\{s_t\} = \text{m.p.e. of } \delta_s \text{ and } y_t, \qquad (20)$$

$$y_t = \{0_t\}\,\sigma_0 + \{1_t\}\,\sigma_1 + \{2_t\}\,\sigma_2 + \dots + \{l_t\}\,\sigma_l. \qquad (21)$$

(iii.) The above relations can be expressed diagrammatically, thus :—

$$\begin{array}{c|cccc|c}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\delta_3 & \times & \times & \times & \times \ldots & \sigma_3 \\
\delta_2 & \times & \times & \times & \times \ldots & \sigma_2 \\
\delta_1 & \times & \times & \times & \times \ldots & \sigma_1 \\
\delta_0 & \times & \times & \times & \times \ldots & \sigma_0 \\
\hline
 & u_0 & u_1 & u_2 & u_3 \ldots & \\
\hline
 & y_0 & y_1 & y_2 & y_3 \ldots &
\end{array}$$

The crosses represent the ( ) coefficients if they are the coefficients of the δ's in the u's and of the y's in the σ's, and the { } coefficients in the converse case.

(iv.) Similarly, if we write ($r = 0, 1, 2, \ldots l$; $t = 0, 1, 2, \ldots l$)

$$[r_t] \equiv \text{m.p.e. of } y_r \text{ and } \sigma_t, \quad \ldots \ldots \ldots \quad (22)$$

then

$$y_r = [r_0]\,\delta_0 + [r_1]\,\delta_1 + [r_2]\,\delta_2 + \ldots + [r_l]\,\delta_l, \quad \ldots \ldots \quad (23)$$

$$\sigma_t = [0_t]\,u_0 + [1_t]\,u_1 + [2_t]\,u_2 + \ldots + [l_t]\,u_l. \quad \ldots \ldots \quad (24)$$

5. *Sums as Conjugates of Differences.*—The cases of importance are those in which the δ's are successive differences of the u's. It will be found that in these cases the σ's are l.cc. of successive sums of the y's.

(i.) Let the δ's be the advancing differences of the u's, *i.e.*

$$\delta_0 = u_0,\ \delta_1 = \Delta u_0, \ldots \delta_r = \Delta^r u_0, \ldots .$$

Then the diagram for the ( ) coefficients is

$$\begin{array}{c|cccc|c}
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\delta_3 & 0 & 0 & 0 & 1 \ldots & \sigma_3 \\
\delta_2 & 0 & 0 & 1 & 3 \ldots & \sigma_2 \\
\delta_1 & 0 & 1 & 2 & 3 \ldots & \sigma_1 \\
\delta_0 & 1 & 1 & 1 & 1 \ldots & \sigma_0 \\
\hline
 & u_0 & u_1 & u_2 & u_3 \ldots & \\
\hline
 & y_0 & y_1 & y_2 & y_3 \ldots &
\end{array}$$

so that

$$\sigma_0 = y_0+y_1+y_2+y_3+\ldots+y_l,$$

$$\sigma_1 = y_1+2y_2+3y_3+4y_4+\ldots+ly_l,$$

$$\sigma_2 = y_2+3y_3+6y_4+10y_5+\ldots+\tfrac{1}{2}l(l-1)\,y_l,$$

and, generally,

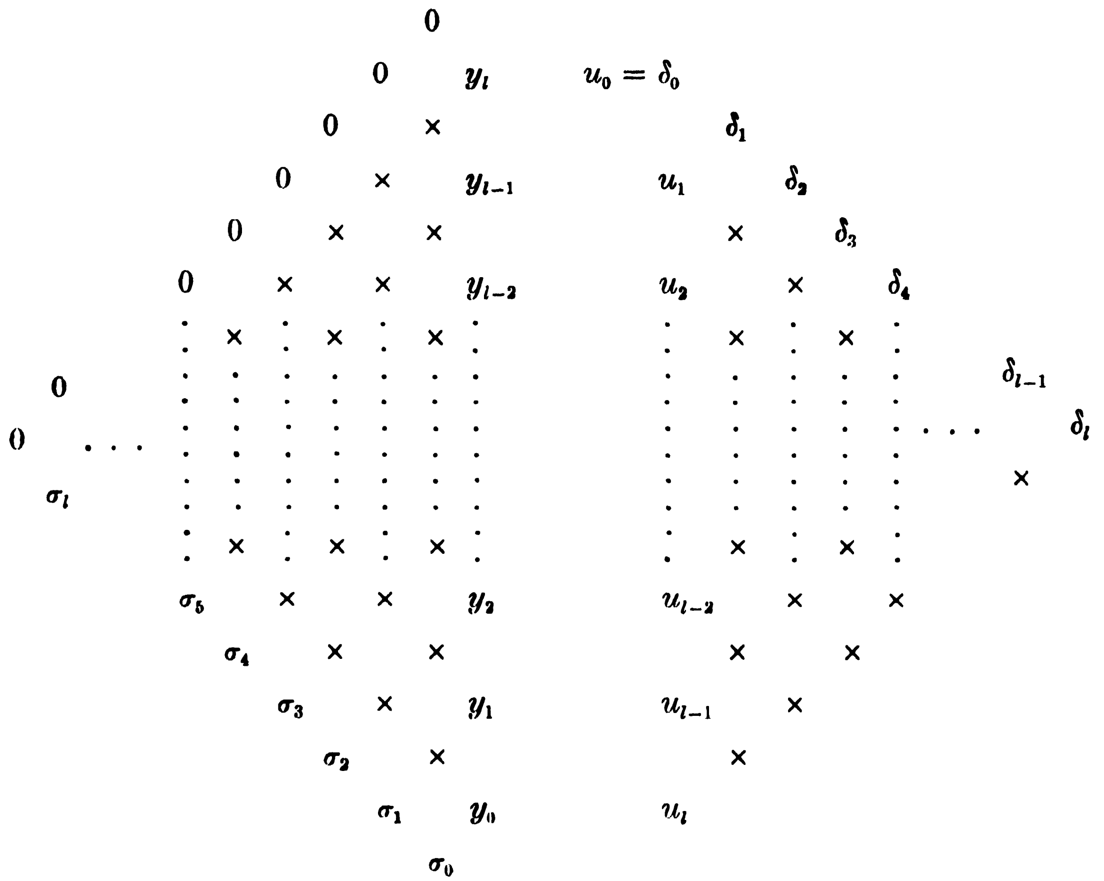
$$\sigma_f = y_f+(f+1,\,1)\,y_{f+1}+(f+2,\,2)\,y_{f+2}+\ldots+(l,\,l-f)\,y_l$$

$$= \sum_{q=f}^{q=l}(q,f)\,y_q; \quad \ldots\ldots\ldots\ldots \quad (25)$$

or, in the notation of "Fitting," § 4, and "Factorial Moments,"

$$\sigma_f = \Sigma''^{f+1}y_f. \quad \ldots\ldots\ldots\ldots \quad (26)$$

The σ's can be obtained by successive summations of the y's in reverse order, *i.e.* from y_l to y_0, as shown in the following diagram, in which the crosses represent entries in a sum- or difference-table :—



(ii.) Let the δ's be the central differences of the u's. Then it will be found in the same way that

(*a*) If the u's are $u_0, u_1, u_2, \ldots u_{2n}$, so that

$$\delta_0 = u_n,\ \delta_1 = \mu\delta u_n,\ \delta_2 = \delta^2 u_n,\ \delta_3 = \mu\delta^3 u_n,\ \ldots,$$

then

$$\sigma_{2h} = \sum_{r=-n}^{r=n} [r, 2h)\, y_{n+r}, \qquad (27)$$

$$\sigma_{2h-1} = \sum_{r=-n}^{r=n} (r, 2h-1]\, y_{n+r}\,; \qquad (28)$$

(*b*) If the u's are $u_0, u_1, u_2, \ldots u_{2n-1}$, so that

$$\delta_0 = \mu u_{n-\frac{1}{2}},\ \delta_1 = \delta u_{n-\frac{1}{2}},\ \delta_2 = \mu\delta^2 u_{n-\frac{1}{2}},\ \delta_3 = \delta^3 u_{n-\frac{1}{2}},\ \ldots,$$

then

$$\sigma_{2h-1} = \sum_{r=-n+1}^{r=n} [r-\tfrac{1}{2}, 2h-1)\, y_{n+r-1}, \qquad (29)$$

$$\sigma_{2h-2} = \sum_{r=-n+1}^{r=n} (r-\tfrac{1}{2}, 2h-2]\, y_{n+r-1}. \qquad (30)$$

(iii.) The values given by (25)–(30) may be expressed in terms of successive sums by the formulæ given in "Factorial Moments." The notation, however, can be simplified. Suppose that we have a set of quantities ... F_0, F_1, F_2, ... corresponding to values ... 0, 1, 2, ... of some variable, and that we form the table of successive differences (and also, if we like, of successive sums) of the F's. Then the Lagrangian formula for F_θ in terms of $F_p, F_{p+1}, \ldots F_{p+t}$, which can be expressed in a good many different ways, may be regarded as the formula for it in terms of the whole (unlimited) set of differences (and sums) which form a triangle with its apex at $\Delta^t F_p$; and we can denote it by $L\{F_\theta\,;\,\Delta^t F_p\}$. With this notation, the above results may be written

(25) $\quad \sigma_f = \left[L\{(-)^f \Sigma^{f+1} y_f\,;\,\Sigma y_t\}\right]_{t=0}^{t=l+1}, \quad . . . . (31)$

(27) $\quad \sigma_{2h} = \left[L\{\mu\sigma^{2h+1} y_n\,;\,\sigma y_{n+t}\}\right]_{t=-n-\frac{1}{2}}^{t=n+\frac{1}{2}}, \quad . . . . (32)$

(28) $\quad \sigma_{2h-1} = \left[L\{-\sigma^{2h} y_n\,;\,\sigma y_{n+t}\}\right]_{t=-n-\frac{1}{2}}^{t=n+\frac{1}{2}}, \quad . . . . (33)$

(29) $\quad \sigma_{2h-1} = \left[L\{-\mu\sigma^{2h} y_{n-\frac{1}{2}}\,;\,\sigma y_{n+t-1}\}\right]_{t=-n+\frac{1}{2}}^{t=n+\frac{1}{2}}, \quad . . . . . . (34)$

(30) $\quad \sigma_{2h-2} = \left[L\{\sigma^{2h-1} y_{n-\frac{1}{2}}\,;\,\sigma y_{n+t-1}\}\right]_{t=-n+\frac{1}{2}}^{t=n+\frac{1}{2}}. \quad . . . . (35)$

The L is distributive as regards the first member inside the { } ; *e.g.*, in the case of (31),

$$A\sigma_2 + B\sigma_3 = \left[L\{A\Sigma^3 y_2 - B\Sigma^4 y_3\,;\,\Sigma y_t\}\right]_{t=0}^{t=l+1}.$$

(iv.) More generally, suppose that the δ's are the successive differences of the u's according to any system of differences; by which we mean that δ_s is either a definite difference of the u's of order s (the u's themselves being differences of order 0) or a l.c. of such differences. Then (r_t) of (14) is a polynomial in r of degree t, and σ_t is of the form $\sum_{q=0}^{q=t} \phi_t(q)\, y_q$, where $\phi_t(q)$ is some polynomial in q of degree t. It follows that any l.c. of $\sigma_0, \sigma_1, \sigma_2, \ldots \sigma_t$ is also of this form.

(v.) If we denote the f^{th} moment of the y's by M_f, then M_f is of the form $\sum_{q=0}^{q=m} \phi_f(q)\, y_q$. Hence σ_t is a l.c. of $M_0, M_1, M_2, \ldots M_t$; and M_t is a l.c. of $\sigma_0, \sigma_1, \sigma_2, \ldots \sigma_t$. More generally, any l.c. of $\sigma_0, \sigma_1, \sigma_2, \ldots \sigma_t$ is a l.c. of $M_0, M_1, M_2, \ldots M_t$; and conversely.

Reduction of Error (General).

6. *General Theorems.*—Let $A, B, C, \ldots P, Q, R, \ldots$ be a set of quantities as in § 2, but all of the same kind. If

$$w \equiv aA+bB+cC+\ldots \text{(with or without terms in } P, Q, R, \ldots),$$

$$x = w+pP+qQ+rR+\ldots,$$

where $a, b, c, \ldots$ are fixed and $p, q, r, \ldots$ are arbitrary, and if we choose $p, q, r, \ldots$ so as to make the m.s.e. of x a minimum, the resulting value of x is called the *improved value of* w, *using* P, Q, R, … *as auxiliaries.* The following are general theorems; some are quite elementary, but it is convenient to state them here. The specially important theorems are (III.) and (XIII.). The assumption mentioned under (VI.) should be noted. If strict proofs of (I.) and (II.) are required, the method should be that of Appendix I., § 2.

(I.) *The m.p.e. of* A *and any l.c. of* A, B, C, … *is the same l.c. of the m.pp.e. of* A *and* A, B, C, … [*i.e.*, m.p.e. of A and $aA+bB+cC+\ldots$ is $a(A\,;\,A)+b(A\,;\,B)+c(A\,;\,C)+\ldots$].

(II.) *The m.s.e. of any l.c. of* A, B, C, …, *or the m.p.e. of any two such l.cc., is found by squaring the former or multiplying the latter and replacing squares and products by the corresponding m.ss.e. and m.pp.e.* [*i.e.*, m.s.e. of $aA+bB+cC+\ldots = a^2(A\,;\,A)+2ab(A\,;\,B)+b^2(B\,;\,B)+2ac(A\,;\,C)+2bc(B\,;\,C)+c^2(C\,;\,C)+\ldots$, and similarly for m.p.e. of $aA+bB+cC+\ldots$ and $a'A+b'B+c'C+\ldots$].

(III.) *If the improved value of* A, *using certain auxiliaries, is* $A+\alpha$, *then the m.p.e. of* $A+\alpha$ *and each of the auxiliaries or* α *or any other l.c. of the auxiliaries is zero.* [Let the auxiliaries be $P, Q, R, \ldots$, and let $A+\alpha = A+pP+qQ+rR+\ldots$. Then the m.s.e. of $A+(p+\theta)P+qQ+rR+\ldots (= A+\alpha+\theta P)$ is $(A+\alpha\,;\,A+\alpha)+2\theta(A+\alpha\,;\,P)+\theta^2(P\,;\,P)$. In order that this may be a minimum for $\theta = 0$, $(A+\alpha\,;\,P)$ must be zero. Similarly for $(A+\alpha\,;\,Q)$, $(A+\alpha\,;\,R)$, …. This proves the first part of the theorem; the second then follows from (I.).] Hence

(IV.) *If the improved values of* A *and of* B, *using in each case the same set of auxiliaries, are* $A+\alpha$ *and* $B+\beta$, *the m.pp.e. of* $A+\alpha$ *and* $B+\beta$, *of* $A+\alpha$ *and* B, *and of* A *and* $B+\beta$, *are all equal;* and

(V.) *If the improved value of* A, *using certain auxiliaries, is* A+α, *and that of* B, *using some only of these, is* B+β, *the m.p.e. of* A+α *and* B+β *is equal to that of* A+α *and* B.

(VI.) *If the improved value of* A, *using* P, Q, R, ... *as auxiliaries, is* A+pP+qQ+rR+..., *the values of* p, q, r, ... *are given by a set of linear relations between* p, q, r, ... *and the m.pp.e. of* A *and* P, Q, R, ... [The equations are given by (III.), viz., since $(A+pP+qQ+rR+\ldots\,;P)=0$, &c.,

$$\begin{aligned}(A\,;P)+p\,(P\,;P)+q\,(Q\,;P)+r\,(R\,;P)+\ldots&=0,\\(A\,;Q)+p\,(P\,;Q)+q\,(Q\,;Q)+r\,(R\,;Q)+\ldots&=0,\\(A\,;R)+p\,(P\,;R)+q\,(Q\,;R)+r\,(R\,;R)+\ldots&=0,\\&\text{\&c.}\end{aligned}$$

It is assumed that the determinant

$$\begin{vmatrix}(P\,;P) & (Q\,;P) & (R\,;P) & \ldots\\(P\,;Q) & (Q\,;Q) & (R\,;Q) & \ldots\\(P\,;R) & (Q\,;R) & (R\,;R) & \ldots\\\vdots & \vdots & \vdots & \end{vmatrix}$$

is not zero; *i.e.* (see § 2 (iii.)) that there is a set conjugate to P, Q, R, ...]

(VII.) *For any particular set of auxiliaries there is one and only one improved value of* A. [This follows from the fact that the equations in (VI.), on the assumption there stated, have one and only one solution.] Hence we get the converse of (III.):—

(VIII.) *If* x *is the sum of* w *and a l.c. of* P, Q, R, ..., *and if the m.p.e. of* x *and each of* P, Q, R, ... *is zero, then* x *is the improved value of* w, *using* P, Q, R, ... *as auxiliaries.* As a particular case:—

(IX.) *If the m.p.e. of* w *and each of* P, Q, R, ... *is zero, the improved value of* w, *using* P, Q, R, ... *as auxiliaries, is the same as the original value.*

(X.) *The improved value of* P, *using* P, Q, R, ... *as auxiliaries, is* P−P = 0. [This follows either from (VIII.) or from taking $A=P$ in the equations in (VI.).] Hence

(XI.) *If* w *is a l.c. of* A, B, C, ... P, Q, R, ..., *the improved value of* w, *using* P, Q, R, ... *as auxiliaries, is the same as that of the quantity obtained by adding to* w *any l.c. of* P, Q, R, ...; and, conversely:—

(XII.) *If the improved values of* w *and of* w′, *using in each case the same set of auxiliaries, are identical, then* w *and* w′ *either are identical or differ by a l.c. of the auxiliaries.*

(XIII.) *If the improved values of* A, B, C, ..., *using in each case the same set of* *auxiliaries, are* $\mathrm{A}+\alpha$, $\mathrm{B}+\beta$, $\mathrm{C}+\gamma, \ldots$, *the improved value of any l.c. of* $\mathrm{A}, \mathrm{B}, \mathrm{C}, \ldots$, *using these auxiliaries, is the same l.c. of* $\mathrm{A}+\alpha$, $\mathrm{B}+\beta$, $\mathrm{C}+\gamma, \ldots$. [Let the l.c. of $A, B, C, \ldots$ be $w \equiv aA+bB+cC+\ldots$. We want to prove that $x \equiv a(A+\alpha)+b(B+\beta)+c(C+\gamma)+\ldots$ is its improved value. We can do this in either of two ways:

(i.) By (III.), the m.p.e. of x and each of the auxiliaries $P, Q, R, \ldots$ is zero; and x differs from w by a l.c. of $P, Q, R, \ldots$. Hence, by (VIII.), x is the improved value of w, using $P, Q, R, \ldots$ as auxiliaries.

(ii.) A more direct proof follows from the linearity of the equations mentioned in (VI.). It is not necessary to set out the proof here.]

(XIV.) *If* $\mathrm{A}, \mathrm{B}, \ldots \mathrm{C}, \mathrm{D}, \ldots \mathrm{P}, \mathrm{Q}, \ldots \mathrm{R}, \mathrm{S}, \ldots$ *fall into two classes* $\mathrm{A}, \mathrm{B}, \ldots \mathrm{P}, \mathrm{Q}, \ldots$ *and* $\mathrm{C}, \mathrm{D}, \ldots, \mathrm{R}, \mathrm{S}, \ldots$, *such that the m.p.e. of each member of the one class and each member of the other class is zero, then the improved value of a l.c. of any of the members, using* $\mathrm{P}, \mathrm{Q}, \ldots \mathrm{R}, \mathrm{S}, \ldots$ *as auxiliaries, is to be found by taking the two classes separately, i.e.*, by using $P, Q, \ldots$ as auxiliaries for the terms in $A, B, \ldots P, Q, \ldots$, and $R, S, \ldots$ as auxiliaries for the terms in $C, D, \ldots R, S, \ldots$. [For the m.s.e. of $aA+bB+\ldots+cC+dD+\ldots+pP+qQ+\ldots+rR+sS+\ldots$ is the sum of those of $aA+bB+\ldots+pP+qQ+\ldots$ and $cC+dD+\ldots+rR+sS+\ldots$, since the m.p.e. of these latter is zero; we cannot reduce the m.s.e. of the first of them by adding terms in $R, S, \ldots$, or that of the second of them by adding terms in $P, Q, \ldots$; and the result is therefore the same as if we considered them separately.]

(XV.) *If the improved value of* w, *using* $\mathrm{P}, \mathrm{Q}, \mathrm{R}, \ldots$ *as auxiliaries, is* $\mathrm{x}=\mathrm{w}+\mathrm{pP}+$*terms in* $\mathrm{Q}, \mathrm{R}, \ldots$, *this is also the improved value of* $\mathrm{w}+\mathrm{pP}$, *using* $\mathrm{Q}, \mathrm{R}, \ldots$ *as auxiliaries.* [For x differs from $w+pP$ by terms in $Q, R, \ldots$, and the m.p.e. of x and each of $Q, R, \ldots$ is zero.] This can be stated more generally as follows:—

(XVI.) *If the improved value of* A, *using a set of auxiliaries* S, *is* $\mathrm{A}+\alpha$, *and if we divide* S *into two sets*, S_1 *and* S_2, *and the corresponding parts of* α *are* α_1 *and* α_2, *then* $\mathrm{A}+\alpha$ *is the improved value of* $\mathrm{A}+\alpha_1$, *using* S_2 *as auxiliaries.* [We may take $\ldots P, Q$ to be S_1, and $R, \ldots$ to be S_2. The theorem states that, if the improved value of A, using $\ldots P, Q, R, \ldots$, is $A+\ldots+pP+qQ+rR+\ldots$, this is also the improved value of $A+\ldots+pP+qQ$, using $R, \ldots$.]

(XVII.) The following corollaries of (III.) may be noticed, though we shall not require them. If the improved values of A and of B, using in each case the same set of auxiliaries, are $A+\alpha$ and $B+\beta$, then

$$(1)\ (A+\alpha\,;\ A+\alpha)=(A\,;\ A)-(\alpha\,;\ \alpha)$$

and

$$(2)\ (A+\alpha\,;\ B+\beta)=(A\,;\ B)-(\alpha\,;\ \beta).$$

7. *Notation: and Particular Values.*—(i.) It will now be convenient to adopt a linear arrangement of the quantities we are dealing with, and we therefore replace

$A, B, C, \ldots P, Q, R, \ldots$ by $\delta_0, \delta_1, \delta_2, \ldots \delta_{j+1}, \delta_{j+2}, \ldots \delta_l$. The order is quite arbitrary, so far as any general theorems are concerned; but it will usually be convenient to place the auxiliaries last. If, for instance, we are using all but $j+1$ as auxiliaries, we denote those not so used by $\delta_0, \delta_1, \delta_2, \ldots \delta_j$, and the auxiliaries by $\delta_{j+1}, \delta_{j+2}, \ldots \delta_l$; the improved values are then denoted by $(\ )_j$.

(ii.) We use the following notation:—

$$(\epsilon_f)_j \equiv \text{improved value of } \delta_f, \text{ using } \delta\text{'s after } \delta_j;$$

$$\mathrm{E}_j \equiv (\epsilon_j)_j = \text{improved value of } \delta_j, \text{ using all subsequent } \delta\text{'s};$$

$$(\lambda_{f,g})_j \equiv \text{m.p.e. of } (\epsilon_f)_j \text{ and } (\epsilon_g)_j;$$

$$\Lambda_j \equiv (\lambda_{j,j})_j = \text{m.s.e. of } \mathrm{E}_j.$$

Where there is no doubt as to the δ's used as auxiliaries, $(\epsilon_f)_j$ and $(\lambda_{f,g})_j$ can be replaced by ϵ_f and $\lambda_{f,g}$.

(iii.) By (X.) of §6—

$$(\epsilon_f)_j = 0 \text{ if } f > j; \quad \ldots \ldots \quad (36)$$

$$(\lambda_{f,g})_j = 0 \text{ if } f > j \text{ or } g > j \quad \ldots \ldots \quad (37)$$

(iv.) By (IV.)—

$$\left.\begin{aligned}(\lambda_{f,g})_j &= \text{m.p.e. of } (\epsilon_f)_j \text{ and } \delta_g \\ &= \text{m.p.e. of } \delta_f \text{ and } (\epsilon_g)_j\end{aligned}\right\}; \quad \ldots \ldots \quad (38)$$

$$\Lambda_j = \text{m.p.e. of } \mathrm{E}_j \text{ and } \delta_j \quad \ldots \ldots \quad (39)$$

8. *Improved Values in terms of Conjugates.*—In "Fitting" I have given some formulæ for improved values in terms of sums. These may be regarded as derivable from a general theorem relating to the expression of improved values in terms of members of the conjugate set. The theorem is given by (XIX.) and (XX.) below; (XVIII.) is a particular case.

(i.) Take any one of the δ's as δ_0. By (6)—

$$\sigma_0 = \eta_{0,0}\delta_0 + \eta_{1,0}\delta_1 + \ldots + \eta_{l,0}\delta_l.$$

Hence $\sigma_0/\eta_{0,0}$ differs from δ_0 by a l.c. of the other δ's. Also the m.p.e. of $\sigma_0/\eta_{0,0}$ and each of these other δ's is zero. It follows from (VIII.) of §6 that $\sigma_0/\eta_{0,0}$ is the improved value of δ_0, using the other δ's as auxiliaries. The m.s.e. of this improved value is $\eta_{0,0}/(\eta_{0,0})^2 = 1/\eta_{0,0}$. Hence—

(XVIII.) *The improved value of any member of the original set, taking all the other members as auxiliaries, is the product of the corresponding member of the conjugate set by a constant; this constant being = the m.s.e. of the improved value.*

(ii.) The first part of the more general theorem is:—

(XIX.) *The improved value of any l.c. of a set of quantities, using those after the first* j + 1 *as auxiliaries, is a l.c. of the first* j + 1 *of the conjugate set.*

For, if w is a l.c. of $\delta_0, \delta_1, \delta_2, \dots \delta_l$, we can (see §2 (v.)) express it as a l.c. of $\sigma_0, \sigma_1, \sigma_2, \dots \sigma_j, \delta_{j+1}, \delta_{j+2}, \dots \delta_l$. Let the result be $(\Sigma)+(\Delta)$, where (Σ) is a l.c. of $\sigma_0, \sigma_1, \sigma_2, \dots \sigma_j$, and (Δ) is a l.c. of $\delta_{j+1}, \delta_{j+2}, \dots \delta_l$ Then (Σ) differs from w by a l.c. of these latter δ's, and the m.p.e. of (Σ) and each of these δ's is 0; hence, by (VIII), (Σ) is the improved value of w, using these δ's as auxiliaries.

(iii.) Further—

(XX.) *The coefficients of the σ's in the improved value of the l.c. are the m.pp.e. of this improved value and the corresponding δ's.*

For, if the improved value of w is x, and we write

$$x = b_0\sigma_0 + b_1\sigma_1 + b_2\sigma_2 + \dots + b_j\sigma_j,$$

then, by the condition of conjugacy of the σ's and the δ's,

$$\text{m.p.e. of } x \text{ and } \delta_f = b_f.$$

(iv.) This would give us a solution of the problem of finding the improved value, if we could find the m.pp.e. Ordinarily, w is or can be expressed in terms of the δ's, and we do not find its improved value independently, but deduce it from those of the δ's up to δ_j. The improved value of δ_h is, by (iii),

$$(\epsilon_h)_j = (\lambda_{0,h})_j\sigma_0 + (\lambda_{1,h})_j\sigma_1 + \dots + (\lambda_{j,h})_j\sigma_j\,; \quad . \; . \; . \; . \; . \; . \; . \quad (40)$$

and the m.pp.e. that we really require are therefore the values of $(\lambda_{f,h})_j$. With regard to this, see §9.

(v.) As the converse of (XIX.) it may be noted that—

(XXI.) *A quantity of the conjugate set, or a l.c. of such quantities, cannot be improved by means of the non-corresponding quantities of the original set; e.g.*, a l.c. of σ_3 and σ_4 cannot be improved by using the δ's, other than δ_3 and δ_4, as auxiliaries. This follows from (IX.) of §6, since the m.p.e. of δ_r and σ_s is 0 unless $r = s$.

(vi.) If, as in §5, there are related conjugate sets of u's and y's, and the δ's are the differences of the u's, it follows from §5 (v.) that (Σ) in (ii.) above is a l.c. of the moments of the y's up to the j^{th}. (XIX.) is therefore a generalisation of the theorem, for a self-conjugate set, that the improved values are l.cc. of the moments; and, in fact, it explains the appearance of the moments in this connexion.

9. *Mean Products of Error of Improved Values.*—(i.) We have found, in §8 (iv.), that

$$(\epsilon_h)_j = (\lambda_{0,h})_j\sigma_0 + (\lambda_{1,h})_j\sigma_1 + \dots + (\lambda_{j,h})_j\sigma_j.$$

To obtain the λ's, we introduce the condition that this shall differ from δ_h by a l.c. of δ's after δ_j.

(ii.) Substituting for σ_0, σ_1, σ_2, ... from (6), this condition gives

$$\begin{aligned}
&(\lambda_{0,h})_j\,(\eta_{0,0}\delta_0+\eta_{1,0}\delta_1+\ldots+\eta_{l,0}\delta_l)\\
&+(\lambda_{1,h})_j\,(\eta_{0,1}\delta_0+\eta_{1,1}\delta_1+\ldots+\eta_{l,1}\delta_l)\\
&+(\lambda_{2,h})_j\,(\eta_{0,2}\delta_0+\eta_{1,2}\delta_1+\ldots+\eta_{l,2}\delta_l)\\
&+\ldots\\
&+(\lambda_{j,h})_j\,(\eta_{0,j}\delta_0+\eta_{1,j}\delta_1+\ldots+\eta_{l,j}\delta_l)\\
&=\delta_h+\text{ terms in }\delta_{j+1},\ \delta_{j+2},\ \ldots\ \delta_l.
\end{aligned}$$

Equating the coefficients of δ_0, δ_1, δ_2, ... δ_j, we find that $(f = 0, 1, 2, \ldots j)$

$$\eta_{f,0}(\lambda_{0,h})_j+\eta_{f,1}(\lambda_{1,h})_j+\ldots+\eta_{f,j}(\lambda_{j,h})_j = 0\ (f\neq h)\text{ or }1\ (f=h). \quad . \quad . \quad . \qquad (41)$$

Let us write

$$\mathrm{H}_j \equiv \begin{vmatrix} \eta_{0,0} & \eta_{0,1} & \eta_{0,2} & \cdots & \eta_{0,j} \\ \eta_{1,0} & \eta_{1,1} & \eta_{1,2} & \cdots & \eta_{1,j} \\ \eta_{2,0} & \eta_{2,1} & \eta_{2,2} & \cdots & \eta_{2,j} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \eta_{j,0} & \eta_{j,1} & \eta_{j,2} & \cdots & \eta_{j,j} \end{vmatrix}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (42)$$

$$\mathrm{H}_{g,h,j} \equiv \text{cofactor of } \eta_{g,h} \text{ in } \mathrm{H}_j = \mathrm{H}_{h,g,j}. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (43)$$

Then

$$\eta_{f,0}\mathrm{H}_{0,h,j}+\eta_{f,1}\mathrm{H}_{1,h,j}+\ldots+\eta_{f,j}\mathrm{H}_{j,h,j} = 0\ (f\neq h)\text{ or }\mathrm{H}_j\ (f=h)\,; \quad . \quad . \quad . \qquad (44)$$

and therefore, by (41),

$$(\lambda_{g,h})_j = \mathrm{H}_{g,h,j}/\mathrm{H}_j. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (45)$$

Substituting in (40), we obtain $(\epsilon_h)_j$ in terms of the σ's.

For the particular case of $g = h = j$, $(\lambda_{g,h})_j$ becomes Λ_j, and $\mathrm{H}_{g,h,j}$ becomes H_{j-1}; so that

$$\Lambda_j = \mathrm{H}_{j-1}/\mathrm{H}_j. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (46)$$

(iii.) As an example, suppose that we have several independent observations, of unequal accuracy, of a single quantity U, and that we wish to obtain a suitably weighted mean, which may be regarded as the improved value of any one of the observations. Let the observed values be u_0, u_1, u_2, ..., and their m.ss.e. a_0^2, a_1^2, a_2^2, ...; the m.pp.e. being 0, since the observations are independent. We take δ_0 to be one of the u's, and δ_1, δ_2, ... to be its successive differences. Then $j = 0$, since the true values of the first and later differences are all 0. Hence, by § 8 (i.), the improved value is $\sigma_0/(\text{m.s.e. of }\sigma_0)$. But, by § 5 (i.), $\sigma_0 = \Sigma y$; and, by § 2 (vi.) (*b*), $y_r = u_r/a_r^2$, so that m.s.e. of $\sigma_0 = \Sigma 1/a^2$. Hence the improved value is $\Sigma(u/a^2)/\Sigma(1/a^2)$.

(iv.) When j is relatively large, the solution given in (ii.) above can only be regarded as a formal one, since it involves calculation of determinants. I have not been able to provide a general solution which shall avoid determinants; but it will be seen in §§ 17–19 that, if we can find the values of certain quantities occurring in the formulæ, we can deduce the λ's and thence the coefficients of the σ's. These latter are important as giving us formulæ which contain only a few terms and are therefore suited for numerical calculation.

10. *Expressions in terms of a Related Set.*—Suppose that there is another set of $l+1$ quantities $u_0, u_1, u_2, \dots u_l$, connected with the δ's by linear relations; and let the set conjugate to the u's be $y_0, y_1, y_2, \dots y_l$. We shall take the relations between the u's and the δ's and between the y's and the δ's to be, as in § 4, ($r = 0, 1, 2, \dots l$)

$$u_r = (r_0)\,\delta_0 + (r_1)\,\delta_1 + (r_2)\,\delta_2 + \dots + (r_l)\,\delta_l, \quad \dots \dots \quad (47)$$

$$y_r = [r_0]\,\delta_0 + [r_1]\,\delta_1 + [r_2]\,\delta_2 + \dots + [r_l]\,\delta_l. \quad \dots \dots \quad (48)$$

(i.) Let the improved value of u_r, using δ's after δ_j, be v_r. Then, from (47), by (XIII.), remembering that, by (36),

$$(\epsilon_f)_j = 0 \text{ if } f > j,$$

we have

$$v_r = (r_0)\,(\epsilon_0)_j + (r_1)\,(\epsilon_1)_j + \dots + (r_l)\,(\epsilon_l)_j \quad \dots \dots \quad (49)$$

$$= (r_0)\,(\epsilon_0)_j + (r_1)\,(\epsilon_1)_j + \dots + (r_j)\,(\epsilon_j)_j. \quad \dots \dots \quad (49\text{A})$$

Thus the v's are related to the ϵ's in the same way that the u's are related to the δ's.

(ii.) Similarly, if the improved value of y_r, using δ's after δ_j, is z_r, we have

$$z_r = [r_0]\,(\epsilon_0)_j + [r_1]\,(\epsilon_1)_j + \dots + [r_l]\,(\epsilon_l)_j \,. \quad \dots \dots \quad (50)$$

$$= [r_0]\,(\epsilon_0)_j + [r_1]\,(\epsilon_1)_j + \dots + [r_j]\,(\epsilon_j)_j\,; \quad \dots \dots \quad (50\text{A})$$

and the z's are related to the ϵ's in the same way that the y's are related to the δ's.

(iii.) Let w be any l.c. of the δ's or of the u's or y's, and let x be its improved value, using δ's after δ_j. Suppose that x is expressed in terms of the y's, the coefficients being $p_0, p_1, p_2, \dots p_l$, so that

$$x = p_0y_0 + p_1y_1 + p_2y_2 + \dots + p_ly_l\,; \quad \dots \dots \quad (51)$$

and let

$$\lambda_g \equiv \text{m.p.e. of } x \text{ and } (\epsilon_g)_j,$$

so that

$$\lambda_g = 0 \text{ if } g > j.$$

Then, by the condition of conjugacy of the u's and the y's,

$$\text{m.p.e. of } x \text{ and } u_r = p_r. \quad \dots \dots \quad (52)$$

Hence, by (IV.) and (49) or (49A),

$$p_r = \text{m.p.e. of } x \text{ and } v_r$$

$$= (r_0)\lambda_0 + (r_1)\lambda_1 + (r_2)\lambda_2 + \ldots + (r_l)\lambda_l \quad \ldots\ldots \quad (53)$$

$$= (r_0)\lambda_0 + (r_1)\lambda_1 + (r_2)\lambda_2 + \ldots + (r_j)\lambda_j. \quad \ldots\ldots \quad (53\text{A})$$

Thus the p's are related to the λ's in the same way that the v's are related to the ϵ's, or the u's to the δ's.

(iv.) Similarly, if

$$x = q_0 u_0 + q_1 u_1 + q_2 u_2 + \ldots + q_l u_l,$$

the q's are related to the λ's in the same way that the z's are related to the ϵ's, or the y's to the δ's.

11. *Special Case of Differences.*—The important practical case is that in which the δ's are successive differences of the u's, in the general sense explained in §5 (iv.). If the differences of order exceeding j are negligible, we can use them as auxiliaries for improving the u's or the δ's or any l.c. of the u's or the δ's.

(i.) Since the δ's are successive differences of the u's, (r_t) is (§5 (iv.)) a polynomial in r of degree t.

(ii.) By §10 (i.) the ϵ's are the differences of the v's according to the same system; and v_r is a polynomial of degree j in r, the differences of the v's of order exceeding j being zero.

(iii.) With the notation of §10 (iii.), the λ's are the differences of the p's according to the same system; and p_r is a polynomial of degree j in r, the differences of the p's of order exceeding j being zero.

(iv.) If we form the differences of the u's in the usual way, there will be l differences of order 1, $l-1$ of order 2, and so on. The $l-j+1$ of order j, namely $\Delta^j u_0, \Delta^j u_1, \ldots \Delta^j u_{l-j}$, will differ from one another by l.cc. of the differences of higher order; and therefore, by (XI.), they will have the same improved value. If we denote this by E, then, if $w \equiv p\Delta^j u_0 + q\Delta^j u_1 + r\Delta^j u_2 + \ldots$, the improved value of w is $(p+q+r+\ldots)$ E.

RELATION OF "REDUCTION OF ERROR" TO "FITTING" (OF A POLYNOMIAL).

12. *Standard System.*—In the case of a standard system, the process of reduction of error and the process of fitting a polynomial (by least squares or by moments) give the same result. The following is a proof of this, not involving the properties of conjugate sets. The observed values are taken to be $u_0, u_1, u_2, \ldots u_l$; and $\sum_t$ denotes summation for $t = 0, 1, 2, \ldots l$.

(i.) If the polynomial which we are fitting to the u's is

$$v_q = a_0 + a_1 q + a_2 q^2 + \ldots + a_j q^j, \quad \ldots\ldots\ldots \quad (54)$$

the values of the a's when we fit by least squares are given ("Fitting," §§1, 2) by the equations ($f = 0, 1, 2, \dots j$)

$$\sum_q q^f . a_0 + \sum_q q^{f+1} . a_1 + \dots + \sum_q q^{f+j} . a_j = \sum_q q^f u_q \equiv \mathrm{M}_f . \quad \dots \dots \quad (55)$$

These are the same equations that are given by the method of moments.

(ii.) The above equation (55) is a statement that the f^{th} moment of the v's is equal to that of the u's. In order to prove that the process of reduction of error, using differences of order exceeding j as auxiliaries, gives the same result, it is sufficient to show (α) that the improved value of u_q as given by this process is of the form of v_q in (54), and (β) that the f^{th} moment of the improved values of the u's is equal to that of the original values for $f = 0, 1, 2, \dots j$.

(iii.) We have shown, in §11 (ii.), that the improved value of u_q is a polynomial of degree j in q. This establishes (α).

(iv.) By (XIII.), the f^{th} moment of the improved values of the u's is equal to the improved value of their f^{th} moment. In order to show that this is equal to the original value of the f^{th} moment, it is sufficient, by (IX.), to show that the m.p.e. of the original f^{th} moment and every difference of order exceeding j is zero.

(v.) Let the k^{th} difference of u_{r-k} be

$$\delta_k \equiv k_0 u_r - k_1 u_{r-1} + \dots + (-)^k k_k u_{r-k}.$$

Then the f^{th} moment is

$$\dots + r^f u_r + (r-1)^f u_{r-1} + (r-2)^f u_{r-2} + \dots,$$

and the m.p.e. of this and δ_k is

$$k_0 r^f - k_1 (r-1)^f + k_2 (r-2)^f - \dots.$$

But this is the k^{th} difference of $(r-k)^f$, and is $= 0$ if $k > f$.

This proves the proposition.

13. *Fitting by Least Squares.*—Next suppose that the set is not self-conjugate. If the δ's were the differences of a set of u's, we should fit a polynomial of degree (say) j to the u's. This suggests that, in the more general case, the u's being connected with the δ's, as in § 10, by the relation ($r = 0, 1, 2, \dots l$)

$$u_r = (r_0)\,\delta_0 + (r_1)\,\delta_1 + (r_2)\,\delta_2 + \dots + (r_l)\,\delta_l, \quad \dots \dots \quad (56)$$

we should try to fit an expression of the form

$$v_r = (r_0)\,\epsilon_0 + (r_1)\,\epsilon_1 + (r_2)\,\epsilon_2 + \dots + (r_j)\,\epsilon_j \quad \dots \dots \quad (57)$$

to the u's by an appropriate method of least squares; the (r)'s being the same as in (56), and the ϵ's being the quantities to be determined.

(i.) If the y's are conjugate to the u's, and if

$$\psi_{r,s} \equiv \text{m.p.e. of } y_r \text{ and } y_s = \psi_{s,r}, \quad . \; . \; . \; . \; . \; . \; . \; . \quad (58)$$

then (see Appendix II.) the direct (or *a priori*) probability of the occurrence of the given set of u's, if the v's as given by (57) were the true U's, is proportional to

$$\exp\left\{-\tfrac{1}{2}\sum_r\sum_s \psi_{r,s}(u_r-v_r)(u_s-v_s)\right\},$$

where $\sum_t$ denotes summation for $t = 0, 1, 2, \ldots l$. The principle of the method of least squares therefore leads us* to choose the ϵ's so as to make

$$\sum_r\sum_s \psi_{r,s}(u_r-v_r)(u_s-v_s)$$

a minimum. Differentiating with regard to each of the ϵ's, this gives ($f = 0, 1, 2, \ldots j$)

$$\sum_s \{(0_f)\psi_{0,s}+(1_f)\psi_{1,s}+\ldots+(l_f)\psi_{l,s}\}(v_s-u_s) = 0. \quad . \; . \; . \; . \quad (59)$$

But, by (58) and (16),

$$\begin{aligned}(0_f)\psi_{0,s}+(1_f)\psi_{1,s}+\ldots+(l_f)\psi_{l,s} &= \text{m.p.e. of } y_s \text{ and } (0_f)y_0+(1_f)y_1+\ldots+(l_f)y_l \\ &= \text{m.p.e. of } y_s \text{ and } \sigma_f. \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad (60)\end{aligned}$$

Denoting this, as in § 4 (iv.), by $[s_f]$, the equations (59) become ($f = 0, 1, 2, \ldots j$)

$$\sum_s [s_f](v_s-u_s) = 0. \quad . \; . \; . \; . \; . \; . \; . \; . \; . \; . \quad (61)$$

(ii.) Instead of fitting an expression of the form given by (57) to the u's we might fit a corresponding expression to the y's. Since

$$y_s = [s_0]\delta_0+[s_1]\delta_1+[s_2]\delta_2+\ldots+[s_l]\delta_l, \quad . \; . \; . \; . \; . \; . \quad (62)$$

the expression to be fitted would be of the form

$$z_s \equiv [s_0]\epsilon_0+[s_1]\epsilon_1+[s_2]\epsilon_2+\ldots+[s_j]\epsilon_j. \quad . \; . \; . \; . \; . \; . \; . \; . \quad (63)$$

---

* Strictly, we ought to choose the ϵ's so as to make $B \, . \, C \exp -\frac{1}{2}P$ a maximum: where

$$P \equiv \sum\sum \psi_{r,s}(u_r-v_r)(u_s-v_s);$$

B is the direct probability of occurrence of the particular ϵ's denoted by $\epsilon_0, \epsilon_1, \epsilon_2, \ldots \epsilon_j$, and is therefore some function of these latter; and $C\exp -\frac{1}{2}P$ is the direct probability of occurrence of the particular values of $u-v$ on the assumption that these values of the ϵ's are the correct ones, C being some function of these ϵ's. But I have assumed, as is commonly done, that the range of practically possible values of the ϵ's is so small that B and C may be treated as constants, so that we have only to consider the maximum value of $\exp -\frac{1}{2}P$.

We should have to choose the ϵ's so as to make

$$\sum_r \sum_s \pi_{r,s} (y_r - z_r)(y_s - z_s)$$

a minimum, where

$$\pi_{r,s} \equiv \text{m.p.e. of } u_r \text{ and } u_s. \quad \ldots \quad (64)$$

This would give

$$\sum_r (r_f)(z_r - y_r) = 0. \quad \ldots \quad (65)$$

(iii.) The ϵ's given by (65) are the same as are given by (61). For we have seen in § 4 that

$$(r_f) = \text{m.p.e. of } u_r \text{ and } \sigma_f, \quad \ldots \quad (66)$$

$$[s_f] = \text{m.p.e. of } y_s \text{ and } \sigma_f. \quad \ldots \quad (67)$$

If we express the u's in terms of the δ's, and write

$$\sum_s [s_f] u_s = \sum_t A_t \delta_t,$$

then, by the condition of conjugacy of the σ's and the δ's, and by (66),

$$\begin{aligned} A_t &= \text{m.p.e. of } \sigma_t \text{ and } \sum_s [s_f] u_s \\ &= \sum_s [s_f](s_t). \end{aligned}$$

This is symmetrical, and we should get the same expression for the coefficient of δ_t in $\sum_r (r_f) y_r$, so that

$$\sum_s [s_f] u_s = \sum_r (r_f) y_r. \quad \ldots \quad (68)$$

Similarly, if we substitute the values of v_s from (57) and of z_r from (63) in $\sum_s [s_f] v_s$ and in $\sum_r (r_f) z_r$, the coefficients of ϵ_t in the resulting expressions are equal. Hence (61) and (65) are identical equations in the ϵ's.

(iv.) The values of the ϵ's as given by these equations are in fact independent of the u's or the y's. For the value of A_t as found in (iii.) above is

$$\begin{aligned} \sum_s [s_f](s_t) &= \text{m.p.e. of } \sum_s [s_f] u_s \text{ and } \sum_s (s_t) y_s \\ &= \text{m.p.e. of } \sigma_f \text{ and } \sigma_t, \quad \ldots \quad (69) \end{aligned}$$

by (24) and (16). Hence, denoting the m.p.e. of σ_f and σ_t, as in § 3, by $\eta_{f,t}$, the ϵ's given by (61) or (65) are the same as would be given by ($f = 0, 1, 2, \ldots j$)

$$\sum_{t=0}^{t=j} \eta_{f,t} \epsilon_t = \sum_{t=0}^{t=l} \eta_{f,t} \delta_t. \quad \ldots \quad (70)$$

(v.) The ordinary method of least squares would consist of making $\sum_s (v_s - u_s)^2$ a minimum, and would lead to equations

$$\sum_s (s_f)(v_s - u_s) = 0,$$

which would not give the most probable values of the ϵ's.

14. *Fitting by Moments.*—(i.) The ordinary method of moments, adapted to the case in which the δ's are not necessarily the successive differences of the u's, would consist in equating the values of $\sum_s (s_f) v_s$ and $\sum_s (s_f) u_s$. This, as will be seen from § 13 (v.), would not give the most probable values of the ϵ's.

(ii.) In order to obtain the most probable values of the ϵ's by equating moments of the v's and of the u's, we must write (say)

$$M_f \equiv \sum_s [s_f] u_s, \quad \ldots \ldots \ldots \quad (71)$$

and define the f^{th} moment of the u's as being M_f or a definite l.c. of M_f, M_{f-1}, $M_{f-2}, \ldots M_0$. But the coefficient of u_s in M_f would then not be given definitely by the relations between the u's and the δ's, but would depend also on the law of correlation of errors of the u's. We see, however, from § 13 (iii.), that we have also

$$M_f = \sum_r (r_f) y_r, \quad \ldots \ldots \ldots \quad (72)$$

and that we get the same result by equating moments, defined in this way, of the y's and the z's. In the ordinary case in which the δ's are successive differences of the u's, the coefficients of the y's in (72) are binomial coefficients, and the ordinary moments fall within the definition given above. It follows that *in fitting a polynomial to a set of quantities* (not being a self-conjugate set) *by the method of moments, the moments which ought to be equated are not those of the quantities themselves and their assumed values, but those of the conjugate set of the former and the corresponding l.cc. of the latter.*

15. *Reduction of Error.*—If we improve the δ's or the u's by means of the δ's after δ_j, the improved values of these latter are zero, and those of the δ's up to δ_j are obtainable from (XXI.) of § 8, which states that the improved values of the σ's from σ_0 to σ_j are equal to the original values. Using (6), this gives ($f = 0, 1, 2, \ldots j$)

$$\sum_{t=0}^{t=j} \eta_{f,t} (\epsilon_t)_j = \sum_{t=0}^{t=l} \eta_{f,t} \delta_t. \quad \ldots \ldots \ldots \quad (73)$$

Comparing this with (70), we see that the ϵ's given by this process are the same as those given by the process of fitting the expression in (57).

16. *Difference of the Two Processes.*—Although the two processes lead to the

same result, they are essentially different. This is explained in § 22 of "Reduction." The main difference may be expressed as follows:—

(i.) In "fitting" we deal directly with the particular case. We assume that the true values follow a specified law, involving unknown constants, and we deduce values for these constants from the data by the principle of inverse probability.

(ii.) In "reduction of error" we do not use inverse probability, and we only deal incidentally with the particular case. We regard the aggregate of the data as one of an indefinitely great number of possible aggregates from the same true values, and we use a method which will reduce as much as possible the m.s.e. of these possible aggregates.

SOME STEPS IN THE GENERAL SOLUTION.

17. *Preliminary.*—(i.) Our object is to find the improved value of any l.c. of the δ's or the u's, and the m.s.e. of this improved value or the m.p.e. of two improved values. Ordinarily, as already stated in § 8 (iv.), the quantity to be improved would be expressible in terms of the δ's, so that we need consider only the improved values of the δ's, *i.e.*, the ϵ's. There are then four problems before us, viz.: (1) expression of the ϵ's in terms of the δ's; (2) expression in terms of the σ's; (3) expression in terms of the y's; (4) determination of the λ's. For practical purposes (2) is more important than (1) or (3), since there will be fewer coefficients involved.

(ii.) Although it does not seem possible to obtain a general solution, otherwise than by determinants, there are some general propositions that indicate stages in the solution. If, without necessarily finding the complete expressions of the ϵ's in terms of the δ's, we can find for each ϵ the coefficient of the first of the auxiliaries, then it will be seen from § 18 that we can find all the ϵ's if we know the E's, and from § 19 (i.) that we can find all the λ's if we know the Λ's. It follows from (40) that in this latter case we can at once obtain the ϵ's in terms of the σ's.

(iii.) We use the notation of § 7, and we also write

$$-\theta_{f,j} \equiv \text{coefficient of } \delta_j \text{ (as auxiliary) in } (\epsilon_f)_{j-1},$$

so that

$$\theta_{f,j} = 0 \text{ if } f > j, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (74)$$

$$\theta_{j,j} = 1. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (75)$$

It should be observed that $\theta_{f,j}$ is not equal to $\theta_{j,f}$. The θ's may be known directly, or, as is shown in (83), we may be able to obtain them from certain of the λ's.

18. *Formula of Progression.*—The quantities which we want to find are the improved values

$$\text{of } \delta_0, \text{ using } \delta_1, \delta_2, \delta_3, \ldots,$$
$$\text{of } \delta_1 \text{ and } \delta_0, \text{ using } \delta_2, \delta_3, \delta_4, \ldots,$$
$$\text{of } \delta_2, \delta_1, \text{ and } \delta_0, \text{ using } \delta_3, \delta_4, \delta_5, \ldots,$$

and so on. There is a formula connecting these, which makes it unnecessary to deal with more than the first quantity in each row; or, if we deal with them all, the formula can be used for checking the results. (An example is given at the end of § 15 of "Reduction.")

We have

$$(\epsilon_f)_{j-1} = \delta_f - \theta_{f,j}\delta_j + \text{ terms in } \delta_{j+1},\ \delta_{j+2},\ \dots .$$

By (XV.), this is the improved value of $\delta_f - \theta_{f,j}\delta_j$, using δ's after δ_j; and therefore, by (XIII.),

$$(\epsilon_f)_{j-1} = (\epsilon_f)_j - \theta_{f,j}(\epsilon_j)_j = (\epsilon_f)_j - \theta_{f,j}\mathrm{E}_j . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (76)$$

Re-arranging, and replacing j by $j-1$, $j-2$, ... $f+1$, and remembering that, by (75), $\theta_{f,f} = 1$, we have

$$\begin{aligned} (\epsilon_f)_j - (\epsilon_f)_{j-1} &= \theta_{f,j}\ \mathrm{E}_j, \\ (\epsilon_f)_{j-1} - (\epsilon_f)_{j-2} &= \theta_{f,j-1}\mathrm{E}_{j-1}, \\ &\ \ \vdots \\ (\epsilon_f)_{f+1} - (\epsilon_f)_f &= \theta_{f,f+1}\mathrm{E}_{f+1}, \\ (\epsilon_f)_f &= \theta_{f,f}\ \mathrm{E}_f. \end{aligned}$$

Hence, by addition,

$$(\epsilon_f)_j = \theta_{f,f}\mathrm{E}_f + \theta_{f,f+1}\mathrm{E}_{f+1} + \theta_{f,f+2}\mathrm{E}_{f+2} + \dots + \theta_{f,j}\mathrm{E}_j . \quad . \quad . \quad . \quad . \quad . \qquad (77)$$

19. *Mean Products of Error (Alternative Formula).*—(i.) By (77) and (38),

$$\begin{aligned} (\lambda_{f,g})_j &= \text{m.p.e. of } \delta_g \text{ and } (\epsilon_f)_j \\ &= \text{m.p.e. of } \delta_g \text{ and } \theta_{f,f}\mathrm{E}_f + \theta_{f,f+1}\mathrm{E}_{f+1} + \dots + \theta_{f,j}\mathrm{E}_j \\ &= \theta_{f,f}(\lambda_{g,f})_f + \theta_{f,f+1}(\lambda_{g,f+1})_{f+1} + \dots + \theta_{f,j}(\lambda_{g,j})_j \\ &= \Sigma\theta_{f,t}(\lambda_{g,t})_t. \end{aligned}$$

The summation has to be made from $t = f$ to $t = j$. But, if $g > f$, we see from (37) that it is sufficient to make the summation from $t = g$. Hence, using "$t = f, g$" to denote summation from $t = f$ or from $t = g$, according as f or g is the greater, we have

$$(\lambda_{f,g})_j = \sum_{t=f,\,g}^{t=j} \theta_{f,t}(\lambda_{g,t})_t . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (78)$$

But, by putting $g = j$ in (78) (or taking the m.p.e. of δ_j and each member of (76)) and then replacing f and j by g and t,

$$(\lambda_{g,t})_t = \theta_{g,t}\Lambda_t . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (79)$$

Hence, substituting in (78),

$$(\lambda_{f,g})_j = \sum_{t=f,\,g}^{t=j} \theta_{f,t}\theta_{g,t}\Lambda_t . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (80)$$

If we can obtain the θ's and the Λ's in a simple form, we thus have a workable formula for calculating the λ's, and thence, by (40), for determining the ϵ's in terms of the σ's.

(ii.) From (80), using (II.) of § 6, we get the m.ss.e. and m.pp.e. of the improved values of any l.cc. of the δ's. Let

$$w \equiv b_0\delta_0 + b_1\delta_1 + \dots + b_l\delta_l, \qquad w' \equiv c_0\delta_0 + c_1\delta_1 + \dots + c_l\delta_l,$$

and let the improved values of w and w', using δ's after δ_j, be x and x'. Then

$$\begin{aligned} \text{m.p.e. of } x \text{ and } x' &= \sum_{t=0}^{t=j} (b_0\theta_{0,t} + b_1\theta_{1,t} + \dots)\,(c_0\theta_{0,t} + c_1\theta_{1,t} + \dots)\,\Lambda_t \\ &= \sum_{t=0}^{t=j} (b_0\theta_{0,t} + b_1\theta_{1,t} + \dots + b_t\theta_{t,t})\,(c_0\theta_{0,t} + c_1\theta_{1,t} + \dots + c_t\theta_{t,t})\,\Lambda_t, \qquad (81) \end{aligned}$$

$$\text{m.s.e. of } x = \sum_{t=0}^{t=j} (b_0\theta_{0,t} + b_1\theta_{1,t} + \dots + b_t\theta_{t,t})^2\Lambda_t. \quad . \quad . \quad . \quad . \quad . \qquad (82)$$

(iii.) We have assumed that the θ's are known. If they are not known directly, but the values of $(\lambda_{f,t})_t$ are known, then, by (79),

$$\theta_{f,t} = (\lambda_{f,t})_t/\Lambda_t. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (83)$$

Substituting in (80),

$$(\lambda_{f,g})_j = \sum_{t=f,g}^{t=j} (\lambda_{f,t})_t\,(\lambda_{g,t})_t/\Lambda_t. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (84)$$

Also (77) is replaced by

$$(\epsilon_f)_j = \sum_{t=f}^{t=j} (\lambda_{f,t})_t\mathrm{E}_t/\Lambda_t. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (85)$$

APPLICATION TO SELF-CONJUGATE SET.

20. *Preliminary.*—(i.) We have now to apply the preceding results to the case in which the u's are a self-conjugate set, so that $(u_r\,;\,u_s) = 0$ $(r \neq s)$, $(u_r\,;\,u_r) = 1$, $y_r = u_r$. We take the δ's to be successive differences of the u's, commencing with a difference of order 0. The δ's to be used as auxiliaries will be those following δ_j; the $(\quad)_j$ will usually be omitted. We shall take the number of u's or of δ's to be m, so that $m = l+1$.

(ii.) There will be three cases to be considered; advancing differences, and central differences for m odd and for m even. For advancing differences the u's will be taken to be $u_0, u_1, \dots u_{m-1}$. For central differences we shall write $m = 2n+1$ or $m = 2n$; and the u's will be $u_{-n}, u_{-n+1}, \dots u_n$ and $u_{-n+1}, u_{-n+2}, \dots u_n$ respectively. We shall require the following m.pp.e., which can be obtained from ordinary difference formulæ.

m.p.e. of $\Delta^f u_0$ and $\Delta^g u_0 = (-)^{f-g}(f+g, f)$, ... (86)

,, $\delta^{2f} u_0$,, $\delta^{2g} u_0 = (-)^{f-g}(2f+2g, f+g)$, ... (87)

,, $\delta^{2f} u_0$,, $\mu\delta^{2g-1} u_0 = 0$, ... (88)

,, $\mu\delta^{2f-1} u_0$,, $\mu\delta^{2g-1} u_0 = (-)^{f-g}(2f+2g-2, f+g-1)/(2f+2g)$, . (89)

,, $\delta^{2f-1} u_{\frac{1}{2}}$,, $\delta^{2g-1} u_{\frac{1}{2}} = (-)^{f-g}(2f+2g-2, f+g-1)$, ... (90)

,, $\delta^{2f-1} u_{\frac{1}{2}}$,, $\mu\delta^{2g-2} u_{\frac{1}{2}} = 0$, ... (91)

,, $\mu\delta^{2f-2} u_{\frac{1}{2}}$,, $\mu\delta^{2g-2} u_{\frac{1}{2}} = (-)^{f-g}(2f+2g-4, f+g-2)/(2f+2g-2)$. (92)

(iii.) For advancing differences we shall have

$$\delta^f \equiv \Delta^f u_0, \qquad \epsilon_f \equiv \Delta^f v_0.$$

The formulæ will be marked (A).

(iv.) For central differences the two cases of m odd and m even must be considered separately; but it will be found that, when the formulæ relating to v_0, $\delta^2 v_0, \ldots$ (m odd) and to $\delta v_{\frac{1}{2}}$, $\delta^3 v_{\frac{1}{2}}, \ldots$ (m even) are properly expressed, they are practically identical in form, as also are those relating to $\mu\delta v_0$, $\mu\delta^3 v_0 \ldots$ (m odd) and to $\mu v_{\frac{1}{2}}$, $\mu\delta^2 v_{\frac{1}{2}}, \ldots$ (m even); and the latter correspond to the former with certain interchanges of ( ] and [ ). We therefore, for $\mu\delta v_0$, $\mu\delta^3 v_0, \ldots$ and $\mu v_{\frac{1}{2}}$, $\mu\delta^2 v_{\frac{1}{2}}, \ldots$, replace θ, E, Λ, λ, by ϕ, I, M, μ, with the appropriate suffixes.

(v.) For $m = 2n+1$ it will be seen from (88), taken with (XIV.) of § 6, that the differences of even and of odd order can be treated independently. The δ's will be u_0, $\delta^2 u_0, \ldots \delta^{2n} u_0$ in the one case and $\mu\delta u_0$, $\mu\delta^3 u_0, \ldots \mu\delta^{2n-1} u_0$ in the other. We shall denote these by δ_0, $\delta_2, \ldots \delta_{2n}$ and δ_1, $\delta_3, \ldots \delta_{2n-1}$ respectively, and shall take j to be $2k$ or $2k+1$ for the former and $2k-1$ or $2k$ for the latter. The subscripts of the θ's and the ϕ's will be modified accordingly; *i.e.*, $\theta_{2f,2k}$ will mean the coefficient of $-\delta_{2k}$ in $(\epsilon_{2f})_{2k-2}$, and similarly for $\phi_{2f-1,2k-1}$. The formulæ for the two cases will be marked (B) and (C) respectively.

(vi.) Similarly for $m = 2n$ we see from (91) that differences of odd and of even order can be treated separately. The δ's are δ_1, $\delta_3, \ldots \delta_{2n-1}$ in the one case, and δ_0, $\delta_2, \ldots \delta_{2n-2}$ in the other, where $\delta_{2f-1} \equiv \delta^{2f-1} u_{\frac{1}{2}}$, $\delta_{2f-2} \equiv \mu\delta^{2f-2} u_{\frac{1}{2}}$; and j is taken to be $2k-1$ or $2k$ for the former and $2k-2$ or $2k-1$ for the latter. Also $\theta_{2f-1,2k-1}$ means the coefficient of $-\delta_{2k-1}$ in $(\epsilon_{2f-1})_{2k-3}$; and similarly for $\phi_{2f-2,2k-2}$. The formulæ will be marked (D) and (E).

(vii.) Writing

$$F\{\alpha,\ \beta,\ \ldots\ ;\ \rho,\ \psi,\ \ldots\} \equiv 1+\frac{\alpha\,.\,\beta\ldots}{\rho\,.\,\psi\ldots}+\frac{\alpha(\alpha+1)\,.\,\beta(\beta+1)\ldots}{\rho(\rho+1)\,.\,\psi(\psi+1)\ldots}+\ldots,$$

where α is a negative integer, it should be noted that

$$F\{-n,\ \beta\ ;\ 1,\ \psi\} = \frac{[\psi-\beta,\ n]}{[\psi,\ n]}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (93)$$

and that, if

$$\alpha = -n, \quad \psi+\chi = -n+\beta+\gamma+1,$$

then*

$$F\{-n,\ \beta,\ \gamma\ ;\ 1,\ \psi,\ \chi\} = \frac{[\psi-\beta,\ n][\chi-\beta,\ n]}{[\psi,\ n][\chi,\ n]} = \frac{[\psi-\gamma,\ n][\chi-\gamma,\ n]}{[\psi,\ n][\chi,\ n]}. \quad . \qquad (94)$$

(viii.) For the central-difference formulæ it will be convenient to write, if r and s are both even or both odd and $s \geqslant r$,

$$\{s,\ r\} \equiv (\tfrac{1}{2})^s\,\frac{(s+r,\ s)(s,\ \tfrac{1}{2}s+\tfrac{1}{2}r)}{r+1}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (95)$$

$$\{r,\ s\} \equiv (\tfrac{1}{2})^r\,\frac{(s+r,\ r)}{(r+1)(\tfrac{1}{2}r+\tfrac{1}{2}s,\ r)}\ ; \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (96)$$

so that, if $k \geqslant f$,

$$\{2k+1,\ 2f+1\} = \frac{[f+\tfrac{3}{2},\ k](k,f)}{2f+2} = \frac{[f+\tfrac{1}{2},\ k+1](k+1,f+1)}{2f+1}, \quad . \quad . \qquad (97)$$

$$\{2k,\ 2f\} = \frac{[f+\tfrac{1}{2},\ k](k,f)}{2f+1} = \frac{[f+\tfrac{3}{2},\ k-1](k-1,f-1)}{2f}, \quad . \quad . \qquad (98)$$

and, if $k \leqslant s$,

$$\{2k+1,\ 2s+1\} = \frac{[s+\tfrac{3}{2},\ k]}{(2k+2)(s,\ k)} = \frac{[s+\tfrac{1}{2},\ k+1]}{(2s+1)(s,\ k)}, \quad . \quad . \quad . \quad . \qquad (99)$$

$$\{2k,\ 2s\} = \frac{[s+\tfrac{1}{2},\ k]}{(2k+1)(s,\ k)}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (100)$$

(ix.) The successive steps are as follows. The formulæ for the ϵ's (the improved values of the differences) in terms of the differences have already been found in "Reduction" and "Fitting"; they depend on certain theorems as to the coefficients when l.cc. of moments or sums are expressed in terms of differences. From these formulæ we get the θ's, and also the E's; and thence we get, in each case, the progression formula supplied by (77). This formula is not really necessary, but it is useful for checking. The Λ's, *i.e.* the m.ss.e. of the E's, are found from (86)–(92), using (IV.) of § 6; this results in certain hypergeometric series, to which we apply (93) and (94). We then get expressions for the λ's, by (80). From these, by (40), we have the values of the ϵ's in terms of the σ's; and also, by § 10 (iii.), the values in terms of the u's (which are identical with the y's). This completes the investigation of the improved values; but we also want to see the extent of the improvement. A further section therefore gives the ratio of the m.s.e. of the improved value to the m.s.e. of the original value, in a form convenient for calculation.

---

* 'Proceedings of the London Mathematical Society,' 2nd series, x., 474.

(x.) The m.p.e. of $\Delta^f u_r$ and $\Delta^f u_{r+t}$ is $(-)^t\,(2f, f+t)$. Hence, in finding the general solution of our problem for the case of a self-conjugate set, we are also finding it for the case of a set in which the m.p.e. of u_r and u_s is of the form $(-)^{sr}\,(2f, f+s-r)$, f being some positive integer; for we can treat these u's as the f^{th} differences of members of a self-conjugate set.

21. *Improved Values.*—(i.) Adapting § 11 (iii.) to the case in which the y's are identical with the u's, we see that, if w is any l.c. of the u's, its improved value, using differences of order exceeding j, is of the form $\Sigma p_r u_r$, where p_r is a polynomial of degree j in r. The improved values of the u's and their differences are obtained from this by means of certain formulæ, given in §§ 6 and 7 (iv.) of "Factorial Moments," for the expression of $\Sigma p_r u_r$ in terms of differences. The results are given in (12), (18), (19), (26), and (25) of "Fitting." From the first of these, replacing $m+1$ by m, we have $(f = 0, 1, 2, \ldots j)$

$$\text{(A)}\qquad \Delta^f v_0 = \sum_{s=0}^{s=m-1} (s, f)\,(j-s, j-f)\,\frac{(j+f+1, j)}{(j+s+1, j)}\,\frac{(m, s+1)}{(m, f+1)}\,\Delta^s u_0\,; \quad \ldots\ldots \quad (101)$$

and from the other four, altering k to $k-1$ in the last, we have $(f = 0, 1, 2, \ldots k$ in (102) and $1, 2, \ldots k$ in (103)–(105))

$$\text{(B)}\qquad \delta^{2f} v_0 = \sum_{s=0}^{s=n} (k+s, k+f)\,(k-s, k-f)\,\frac{(2k+2f+1, 2k)}{(2k+2s+1, 2k)}\,\frac{[n+\frac{1}{2}, 2s+1)}{[n+\frac{1}{2}, 2f+1)}\,\delta^{2s} u_0\ . \quad (102)$$

$$= \delta^{2f} u_0 + (-)^{k-f} \sum_{s=k+1}^{s=n} \frac{s-k}{s-f}\,\frac{2f+2}{2k+2}\,\frac{\{2k+1, 2f+1\}}{\{2k+1, 2s+1\}}\,\frac{[\frac{1}{2}m, 2s+1)}{[\frac{1}{2}m, 2f+1)}\,\delta^{2s} u_0,\ . \quad (102\text{A})$$

$$\text{(C)}\quad \mu\delta^{2f-1} v_0 = \sum_{s=1}^{s=n} (k+s-1, k+f-1)\,(k-s, k-f)\,\frac{(2k+2f-1, 2k-1)}{(2k+2s-1, 2k-1)}\,\frac{(n+\frac{1}{2}, 2s]}{(n+\frac{1}{2}, 2f]}\,\mu\delta^{2s-1} u_0 \quad \ldots \quad (103)$$

$$= \mu\delta^{2f-1} u_0 + (-)^{k-f} \sum_{s=k+1}^{s=n} \frac{s-k}{s-f}\,\frac{2f+1}{2k+1}\,\frac{\{2k, 2f\}}{\{2k, 2s\}}\,\frac{(\frac{1}{2}m, 2s]}{(\frac{1}{2}m, 2f]}\,\mu\delta^{2s-1} u_0,\ \ldots \quad (103\text{A})$$

$$\text{(D)}\qquad \delta^{2f-1} v_{\frac{1}{2}} = \sum_{s=1}^{s=n} (k+s-1, k+f-1)\,(k-s, k-f)\,\frac{(2k+2f-1, 2k-1)}{(2k+2s-1, 2k-1)}\,\frac{[n, 2s)}{[n, 2f)}\,\delta^{2s-1} u_{\frac{1}{2}} \quad \ldots \quad (104)$$

$$= \delta^{2f-1} u_{\frac{1}{2}} + (-)^{k-f} \sum_{s=k+1}^{s=n} \frac{s-k}{s-f}\,\frac{2f+1}{2k+1}\,\frac{\{2k, 2f\}}{\{2k, 2s\}}\,\frac{[\frac{1}{2}m, 2s)}{[\frac{1}{2}m, 2f)}\,\delta^{2s-1} u_{\frac{1}{2}},\ \ldots \quad (104\text{A})$$

(E) $$\mu\delta^{2f-2}v_{\frac{1}{2}} = \sum_{s=1}^{s=n}(k+s-2,\ k+f-2)(k-s,\ k-f)\frac{(2k+2f-3,\ 2k-2)}{(2k+2s-3,\ 2k-2)}\frac{(n,\ 2s-1]}{(n,\ 2f-1]}\mu\delta^{2s-2}u_{\frac{1}{2}} \quad . \ . \ . \ (105)$$

$$= \mu\delta^{2f-2}u_{\frac{1}{2}} + (-)^{k-f}\sum_{s=k+1}^{s=n}\frac{s-k}{s-f}\frac{2f}{2k}\frac{\{2k-1,\ 2f-1\}}{\{2k-1,\ 2s-1\}}\frac{(\frac{1}{2}m,\ 2s-1]}{(\frac{1}{2}m,\ 2f-1]}\mu\delta^{2s-2}u_{\frac{1}{2}}. \quad . \ . \ . \ (105\text{A})$$

(ii.) From (i.) we obtain

(A) $$\theta_{f,j} = (-)^{j-f}(j,f)\frac{(j+f,\ j-1)}{(2j,\ j-1)}\frac{(m,\ j+1)}{(m,\ f+1)}, \quad . \ . \ . \ . \ (106)$$

(B) $$\theta_{2f,2k} = (-)^{k-f}\frac{\{2k,\ 2f\}}{\{2k,\ 2k\}}\frac{[\frac{1}{2}m,\ 2k+1)}{[\frac{1}{2}m,\ 2f+1)}, \quad . \ . \ . \ . \ . \ . \ (107)$$

(C) $$\phi_{2f-1,2k-1} = (-)^{k-f}\frac{\{2k-1,\ 2f-1\}}{\{2k-1,\ 2k-1\}}\frac{(\frac{1}{2}m,\ 2k]}{(\frac{1}{2}m,\ 2f]}, \quad . \ . \ . \ . \ . \ (108)$$

(D) $$\theta_{2f-1,2k-1} = (-)^{k-f}\frac{\{2k-1,\ 2f-1\}}{\{2k-1,\ 2k-1\}}\frac{[\frac{1}{2}m,\ 2k)}{[\frac{1}{2}m,\ 2f)}, \quad . \ . \ . \ . \ . \ (109)$$

(E) $$\phi_{2f-2,2k-2} = (-)^{k-f}\frac{\{2k-2,\ 2f-2\}}{\{2k-2,\ 2k-2\}}\frac{(\frac{1}{2}m,\ 2k-1]}{(\frac{1}{2}m,\ 2f-1]}. \quad . \ . \ . \ . \ (110)$$

(iii.) Also, by putting $f=j$ in (101) and $f=k$ in (102)–(105),

(A) $$\mathrm{E}_j = \Delta^j v_0 = \sum_{s=j}^{s=m-1}(s,j)\frac{(2j+1,\ j)}{(j+s+1,\ j)}\frac{(m,\ s+1)}{(m,\ j+1)}\Delta^s u_0, \quad . \ . \ . \ (111)$$

(B) $$\mathrm{E}_{2k} = \delta^{2k}v_0 = \sum_{s=k}^{s=n}\frac{\{2k+1,\ 2k+1\}}{\{2k+1,\ 2s+1\}}\frac{[\frac{1}{2}m,\ 2s+1)}{[\frac{1}{2}m,\ 2k+1)}\delta^{2s}u_0, \quad . \ . \ . \ . \ (112)$$

(C) $$\mathrm{I}_{2k-1} = \mu\delta^{2k-1}v_0 = \sum_{s=k}^{s=n}\frac{\{2k,\ 2k\}}{\{2k,\ 2s\}}\frac{(\frac{1}{2}m,\ 2s]}{(\frac{1}{2}m,\ 2k]}\mu\delta^{2s-1}u_0, \quad . \ . \ . \ . \ . \ . \ (113)$$

(D) $$\mathrm{E}_{2k-1} = \delta^{2k-1}v_{\frac{1}{2}} = \sum_{s=k}^{s=n}\frac{\{2k,\ 2k\}}{\{2k,\ 2s\}}\frac{[\frac{1}{2}m,\ 2s)}{[\frac{1}{2}m,\ 2k)}\delta^{2s-1}u_{\frac{1}{2}}, \quad . \ . \ . \ . \ . \ . \ (114)$$

(E) $$\mathrm{I}_{2k-2} = \mu\delta^{2k-2}v_{\frac{1}{2}} = \sum_{s=k}^{s=n}\frac{\{2k-1,\ 2k-1\}}{\{2k-1,\ 2s-1\}}\frac{(\frac{1}{2}m,\ 2s-1]}{(\frac{1}{2}m,\ 2k-1]}\mu\delta^{2s-2}u_{\frac{1}{2}}. \quad . \ . \ . \ (115)$$

(iv.) It has been pointed out in § 11 (iv.) that the differences of order j all have the same improved value. It follows that (112)–(115) are particular cases of (111), expressed in terms of central differences; the proper values being taken for m and for j, and u_0 being altered to u_{-n} for (112) and (113) and to u_{-n+1} for (114) and (115). We can verify this by expressing the E's in terms of the differences of order j. For (A) we have

$$\Delta^s u_0 = \{(1+\Delta)-1\}^{s-j}\Delta^j u_0 = \Delta^j u_{s-j} - (s-j,\ 1)\Delta^j u_{s-j-1} + \dots.$$

Substituting in (111), and rearranging the terms, it will be found that the coefficient of $\Delta^j u_r$ is

$$(j+r,\, j)\frac{(2j+1,\, j)\,(m,\, j+r+1)}{(2j+r+1,\, j)\,(m,\, j+1)}F\{-m+j+r+1,\, j+r+1\,;\, 1,\, 2j+r+2\}$$
$$= (j+r,\, j)\,(m-r-1,\, j)/(m,\, 2j+1]\,;$$

so that

$$\text{(A)} \qquad \mathrm{E}_j = \sum_{r=0}^{r=m-j-1}(j+r,\, j)\frac{(m-r-1,\, j)}{(m,\, 2j+1]}\Delta^j u_r. \qquad (116)$$

Similarly from (112)–(115)

$$\text{(B)} \qquad \mathrm{E}_{2k} = \sum_{r=-n+k}^{r=n-k}\frac{(n+k+r,\, 2k)\,(n+k-r,\, 2k)}{(m,\, 4k+1]}\delta^{2k}u_r, \qquad (117)$$

$$\text{(C)} \qquad \mathrm{I}_{2k-1} = \sum_{r=0}^{r=n-k}\frac{(n+k+r,\, 2k-1)\,(n+k-r-1,\, 2k-1)}{(m,\, 4k-1]}(\delta^{2k-1}u_{-r-\frac{1}{2}}+\delta^{2k-1}u_{r+\frac{1}{2}}), \qquad (118)$$

$$\text{(D)} \qquad \mathrm{E}_{2k-1} = \sum_{r=-n+k}^{r=n-k}\frac{(n+k+r-1,\, 2k-1)\,(n+k-r-1,\, 2k-1)}{(m,\, 4k-1]}\delta^{2k-1}u_{r+\frac{1}{2}}, \qquad (119)$$

$$\text{(E)} \qquad \mathrm{I}_{2k-2} = \sum_{r=0}^{r=n-k}\frac{(n+k+r-1,\, 2k-2)\,(n+k-r-2,\, 2k-2)}{(m,\, 4k-3]}(\delta^{2k-2}u_{-r}+\delta^{2k-2}u_{r+1}). \qquad (120)$$

The identity of (117)–(120) with (116), the u's being altered as explained above, is easily verified.

(v.) The formula of progression (77) takes simple forms if we attach the factors involving m to the δ's; for m then disappears from the formula.

(A) Writing

$$A_t \equiv (m,\, t+1)\,\mathrm{E}_t = (m,\, t+1)\,(\Delta^t v_0)_t,$$

so that, in effect, we take δ_f to be $(m, f+1)\,\Delta^f u_0$, (77) gives

$$\text{(A)} \qquad (m, f+1)\,\Delta^f v_0 = \sum_{t=f}^{t=j}(-)^{t-f}(t, f)\frac{(t+f,\, t-1)}{(2t,\, t-1)}A_t. \qquad (121)$$

For $j = 3$, for instance, we should have

$$\left.\begin{aligned}(m,\, 1)\,v_0 &= A_0 - A_1 + \tfrac{1}{2}A_2 - \tfrac{1}{5}A_3\\ (m,\, 2)\,\Delta v_0 &= \phantom{A_0-{}} A_1 - \tfrac{3}{2}A_2 + \tfrac{6}{5}A_3\\ (m,\, 3)\,\Delta^2 v_0 &= \phantom{A_0-A_1+\tfrac{1}{2}} A_2 - 2A_3\\ (m,\, 4)\,\Delta^3 v_0 &= \phantom{A_0-A_1+\tfrac{1}{2}A_2-\tfrac{1}{5}} A_3\end{aligned}\right\},$$

where A_0 is the value of $(m, 1)\,v_0$ for $j = 0$, A_1 is the value of $(m, 2)\,\Delta v_0$ for $j = 1$, and so on. This may be verified by § 5 (i.)–(iii.) of "Fitting."

(B) (C) Writing

$$P_{2t} \equiv [\tfrac{1}{2}m,\, 2t+1)\,\mathrm{E}_{2t},\ Q_{2t-1} \equiv (\tfrac{1}{2}m,\, 2t]\,\mathrm{I}_{2t-1},$$

we have

$$\text{(B)} \qquad [\tfrac{1}{2}m, 2f+1)\, \delta^{2f} v_0 = \sum_{t=f}^{t=k} (-)^{t-f} \frac{\{2t, 2f\}}{\{2t, 2t\}} P_{2t}, \tag{122}$$

$$\text{(C)} \qquad (\tfrac{1}{2}m, 2f]\, \mu\delta^{2f-1} v_0 = \sum_{t=f}^{t=k} (-)^{t-f} \frac{\{2t-1, 2f-1\}}{\{2t-1, 2t-1\}} Q_{2t-1}. \tag{123}$$

The first of these has been given, for $f = 0$, in "Reduction," § 15 (v.), p. 362; the notation of P_{2t} differing, however, in a factor $\frac{1}{2}m$.

(D) (E) Writing

$$P_{2t-1} \equiv [\tfrac{1}{2}m, 2t)\, \mathrm{E}_{2t-1}, \quad Q_{2t-2} \equiv (\tfrac{1}{2}m, 2t-1]\, \mathrm{I}_{2t-2},$$

we have

$$\text{(D)} \qquad [\tfrac{1}{2}m, 2f)\, \delta^{2f-1} v_{\frac{1}{2}} = \sum_{t=f}^{t=k} (-)^{t-f} \frac{\{2t-1, 2f-1\}}{\{2t-1, 2t-1\}} P_{2t-1}, \tag{124}$$

$$\text{(E)} \qquad (\tfrac{1}{2}m, 2f-1]\, \mu\delta^{2f-2} v_{\frac{1}{2}} = \sum_{t=f}^{t=k} (-)^{t-f} \frac{\{2t-2, 2f-2\}}{\{2t-2, 2t-2\}} Q_{2t-2}. \tag{125}$$

(vi.) The Λ's are obtained from the E's by means of (IV.) of § 6; *e.g.*, for advancing differences, m.s.e. of E_j = m.p.e. of E_j and $\Delta^j u_0$. We can use either the values of the E's given by (111)–(115) or those given by (116)–(120); for the former we require the m.pp.e. given in (86), (87), (89), (90), (92), and for the latter the value of the m.p.e. of two differences of order f as given in § 20 (x.). Using the former method, we get the following results:—

$$\text{(A)} \qquad \begin{aligned} \Lambda_j &= \sum_{s=j}^{s=m-1} (-)^{s-j} (s, j) \frac{(2j+1, j)}{(s+j+1, j)} \frac{(m, s+1)}{(m, j+1)} (s+j, j) \\ &= (2j, j)\, F\{-m+j+1, 2j+1 \,;\, 1, 2j+2\} \\ &= (2j, j)\, (m-j-1)\,!\,(2j+1)\,!/(m+j)\,! \\ &= (2j, j)/(m, 2j+1]\,; \end{aligned} \tag{126}$$

$$\text{(B)} \qquad \begin{aligned} \Lambda_{2k} &= \sum_{s=k}^{s=n} (-)^{s-k} \frac{\{2k+1, 2k+1\}}{\{2k+1, 2s+1\}} \frac{[\frac{1}{2}m, 2s+1)}{[\frac{1}{2}m, 2k+1)} (2s+2k, s+k) \\ &= (4k, 2k)\, F\{-n+k, n+k+1, 2k+\tfrac{1}{2} \,;\, 1, 2k+1, 2k+\tfrac{3}{2}\} \\ &= (4k, 2k)/(m, 4k+1]\,; \end{aligned} \tag{127}$$

$$\text{(C)} \qquad \begin{aligned} \mathrm{M}_{2k-1} &= \sum_{s=k}^{s=n} (-)^{s-k} \frac{\{2k, 2k\}}{\{2k, 2s\}} \frac{(\frac{1}{2}m, 2s]}{(\frac{1}{2}m, 2k]} \frac{(2s+2k-2, s+k-1)}{2s+2k} \\ &= (4k-2, 2k-1)/(4k) \,.\, F\{-n+k, n+k+1, 2k-\tfrac{1}{2} \,;\, 1, 2k+1, 2k+\tfrac{1}{2}\} \\ &= (4k-2, 2k-1)/(m, 4k-1]\,; \end{aligned} \tag{128}$$

$$\text{(D)} \qquad \begin{aligned} \Lambda_{2k-1} &= \sum_{s=k}^{s=n} (-)^{s-k} \frac{\{2k, 2k\}}{\{2k, 2s\}} \frac{[\frac{1}{2}m, 2s)}{[\frac{1}{2}m, 2k)} (2s+2k-2, s+k-1) \\ &= (4k-2, 2k-1)\, F\{-n+k, n+k, 2k-\tfrac{1}{2} \,;\, 1, 2k, 2k+\tfrac{1}{2}\} \\ &= (4k-2, 2k-1)/(m, 4k-1]\,; \end{aligned} \tag{129}$$

$$\text{(E)}\qquad \mathrm{M}_{2k-2} = \sum_{s=k}^{s=n} (-)^{s-k} \frac{\{2k-1,\ 2k-1\}}{\{2k-1,\ 2s-1\}} \frac{(\tfrac{1}{2}m,\ 2s-1]}{(\tfrac{1}{2}m,\ 2k-1]} \frac{(2s+2k-4,\ s+k-2)}{2s+2k-2}.$$

$$= (4k-4,\ 2k-2)/(4k-2)\,.\,F\{-n+k,\ n+k,\ 2k-\tfrac{3}{2}\,;\ 1,\ 2k,\ 2k-\tfrac{1}{2}\}$$

$$= (4k-4,\ 2k-2)/(m,\ 4k-3]. \qquad (130)$$

It would, of course, in view of the identity of the E's as stated in (iv.) above, have been sufficient to obtain (126) and deduce (127)–(130) from it.

(vii.) The Λ's having been found, the λ's, *i.e.*, the m.pp.e. of the improved values of $\Delta^f u_0$ and $\Delta^g u_0$, etc., are obtained by (80). The θ's and ϕ's being as in (ii.) above,

$$\text{(A)}\qquad \lambda_{f,g} = \sum_{t=f,g}^{t=j} \theta_{f,t}\,\theta_{g,t}\,(2t,\ t)/(m,\ 2t+1], \qquad (131)$$

$$\text{(B)}\qquad \lambda_{2f,2g} = \sum_{t=f,g}^{t=k} \theta_{2f,2t}\,\theta_{2g,2t}\,(4t,\ 2t)/(m,\ 4t+1], \qquad (132)$$

$$\text{(C)}\qquad \mu_{2f-1,2g-1} = \sum_{t=f,g}^{t=k} \phi_{2f-1,2t-1}\,\phi_{2g-1,2t-1}\,(4t-2,\ 2t-1)/(m,\ 4t-1], \qquad (133)$$

$$\text{(D)}\qquad \lambda_{2f-1,2g-1} = \sum_{t=f,g}^{t=k} \theta_{2f-1,2t-1}\,\theta_{2g-1,2t-1}\,(4t-2,\ 2t-1)/(m,\ 4t-1], \qquad (134)$$

$$\text{(E)}\qquad \mu_{2f-2,2g-2} = \sum_{t=f,g}^{t=k} \phi_{2f-2,2t-2}\,\phi_{2g-2,2t-2}\,(4t-4,\ 2t-2)/(m,\ 4t-3]. \qquad (135)$$

To find the m.s.e. of the improved value of any l.c. of the differences, or the m.p.e. of two such l.cc., we apply (II.) of § 6, as in § 19 (ii.). Thus, for $m = 2n+1$,

m.p.e. of $b_0 v_0 + b_1\mu\delta v_0 + b_2\delta^2 v_0 + \ldots + b_{2k}\delta^{2k} v_0$ and $c_0 v_0 + c_1\mu\delta v_0 + c_2\delta^2 v_0 + \ldots + c_{2k}\delta^{2k} v_0$

$$= \sum_{t=0}^{t=k} \left\{ \sum_{f=0}^{f=t} b_{2f}\,\theta_{2f,2t} \right\} \left\{ \sum_{g=0}^{g=t} c_{2g}\,\theta_{2g,2t} \right\} (4t,\ 2t)/(m,\ 4t+1]$$

$$+ \sum_{t=1}^{t=k} \left\{ \sum_{f=1}^{f=t} b_{2f-1}\phi_{2f-1,2t-1} \right\} \left\{ \sum_{g=1}^{g=t} c_{2g-1}\phi_{2g-1,2t-1} \right\} (4t-2,\ 2t-1)/(m,\ 4t-1].$$

(viii.) The λ's having been found, the ϵ's are then given in terms of the σ's by (40). Using the expressions for the σ's given in (26) and (31)–(35), we get the following :—

$$\text{(A)}\qquad \Delta^f v_0 = \sum_{g=0}^{g=j} \lambda_{f,g}\,\Sigma''^{g+1} u_g \qquad (136)$$

$$= \left[L\left\{ \sum_{g=0}^{g=j} (-)^g \lambda_{f,g}\,\Sigma^{g+1} u_g\ ;\ \Sigma u_t \right\} \right]_{t=0}^{t=m}; \qquad (136\text{A})$$

$$\text{(B)}\qquad \delta^{2f} v_0 = \left[L\left\{ \sum_{g=0}^{g=k} \lambda_{2f,2g}\,\mu\sigma^{2g+1} u_0\ ;\ \sigma u_t \right\} \right]_{t=-\frac{1}{2}m}^{t=\frac{1}{2}m}; \qquad (137)$$

$$\text{(C)}\qquad \mu\delta^{2f-1} v_0 = \left[L\left\{ -\sum_{g=1}^{g=k} \mu_{2f-1,2g-1}\sigma^{2g} u_0\ ;\ \sigma u_t \right\} \right]_{t=-\frac{1}{2}m}^{t=\frac{1}{2}m}; \qquad (138)$$

$$\text{(D)} \qquad \delta^{2f-1}v_{\frac{1}{2}} = \left[L\left\{-\sum_{g=1}^{g=k} \lambda_{2f-1,2g-1}\mu\sigma^{2g}u_{\frac{1}{2}}\,;\ \sigma u_{t+\frac{1}{2}}\right\}\right]_{t=-\frac{1}{2}m}^{t=\frac{1}{2}m}; \quad \ldots \quad (139)$$

$$\text{(E)} \qquad \mu\delta^{2f-2}v_{\frac{1}{2}} = \left[L\left\{\sum_{g=1}^{g=k} \mu_{2f-2,2g-2}\sigma^{2g-1}u_{\frac{1}{2}}\,;\ \sigma u_{t+\frac{1}{2}}\right\}\right]_{t=-\frac{1}{2}m}^{t=\frac{1}{2}m}. \quad \ldots \quad (140)$$

The expression in (136) differs slightly from that given for $\Delta^f v_0$ in "Fitting," (15) and § 5; see Appendix III.

(ix.) Finally, we want to express the ϵ's in terms of the u's. This is done by means of the general theorem in § 10 (iii.), that the coefficients of the y's in any ϵ are related to the m.pp.e. of this ϵ and the ϵ's in the same way that the v's are related to the ϵ's. Thus we find that—

$$\text{(A)} \qquad \text{If} \quad \Delta^f v_0 = \sum_{r=0}^{r=m} p_r u_r, \quad \text{then} \quad p_r = \sum_{g=0}^{g=f} (r, g)\,\lambda_{f,g}\,; \quad \ldots \quad (141)$$

$$\text{(B)} \qquad \text{If} \quad \delta^{2f} v_0 = \sum_{r=-n}^{r=n} p_r u_r, \quad \text{then} \quad p_r = \sum_{g=0}^{g=k} \left[r, 2g\right)\lambda_{2f,2g}\,; \quad \ldots \quad (142)$$

$$\text{(C)} \qquad \text{If} \quad \mu\delta^{2f-1} v_0 = \sum_{r=-n}^{r=n} p_r u_r, \quad \text{then} \quad p_r = \sum_{g=1}^{g=k} \left(r, 2g-1\right]\mu_{2f-1,2g-1}\,; \quad \ldots \quad (143)$$

$$\text{(D)} \qquad \text{If} \quad \delta^{2f-1} v_{\frac{1}{2}} = \sum_{r=-n+1}^{r=n} p_r u_r, \quad \text{then} \quad p_r = \sum_{g=1}^{g=k} \left[r-\tfrac{1}{2}, 2g-1\right)\lambda_{2f-1,2g-1}\,; \quad \ldots \quad (144)$$

$$\text{(E)} \qquad \text{If} \quad \mu\delta^{2f-2} v_{\frac{1}{2}} = \sum_{r=-n+1}^{r=n} p_r u_r, \quad \text{then} \quad p_r = \sum_{g=1}^{g=k} \left(r-\tfrac{1}{2}, 2g-2\right]\mu_{2f-2,2g-2}. \quad \ldots \quad (145)$$

For a comparison of these formulæ with those given in "Fitting," see Appendix IV.

22. *Extent of Improvement* (*Central Differences*).—A question of practical importance is the extent to which the use of these formulæ actually reduces the m.s.e. of some selected quantity, such as, for the cases marked (B), u_0 or $\delta^{2f}u_0$. The m.ss.e. of the various improved values are found from (131)–(135), by putting $g = f$. Comparing these with the m.ss.e. of the original values, for the central-difference formulæ (which are the important ones for practical use), we obtain the following:—

$$\text{(B)} \quad \frac{\text{m.s.e. of } \delta^{2f}v_0}{\text{m.s.e. of } \delta^{2f}u_0} = \frac{1}{(4f, 2f)}\sum_{t=f}^{t=k}\left\{\frac{\{2t, 2f\}\left[\frac{1}{2}m, 2t+1\right)}{\{2t, 2t\}\left[\frac{1}{2}m, 2f+1\right)}\right\}^2 \frac{(4t, 2t)}{\left(m, 4t+1\right]}$$

$$= \frac{1}{\left(m, 4f+1\right]}\sum_{t=f}^{t=k}\frac{4t+1}{4f+1}\left\{\frac{\{2t, 2f\}}{\{2f, 2f\}}\right\}^2$$

$$\frac{\{m^2-(2f+1)^2\}\{m^2-(2f+3)^2\}\ldots\{m^2-(2t-1)^2\}}{\{m^2-(2f+2)^2\}\{m^2-(2f+4)^2\}\ldots\{m^2-(2t)^2\}}; \quad (146)$$

$$\text{(C)}\quad \frac{\text{m.s.e. of } \mu\delta^{2f-1}v_0}{\text{m.s.e. of } \mu\delta^{2f-1}u_0} = \frac{4f}{(4f-2,\ 2f-1)} \sum_{t=f}^{t=k} \left\{ \frac{\{2t-1, 2f-1\}}{\{2t-1,\ 2t-1\}} \frac{(\tfrac{1}{2}m,\ 2t]}{(\tfrac{1}{2}m, 2f]} \right\}^2 \frac{(4t-2,\ 2t-1)}{(m,\ 4t-1]}$$

$$= \frac{4f}{(m,\ 4f-1]} \sum_{t=f}^{t=k} \frac{4t-1}{4f-1} \left\{ \frac{\{2t-1,\ 2f-1\}}{\{2f-1, 2f-1\}} \right\}^2$$

$$\frac{\{m^2-(2f+1)^2\}\ \{m^2-(2f+3)^2\} \ldots \{m^2-(2t-1)^2\}}{\{m^2-(2f)^2\}\ \{m^2-(2f+2)^2\} \ldots \{m^2-(2t-2)^2\}}; \quad (147)$$

$$\text{(D)}\quad \frac{\text{m.s.e. of } \delta^{2f-1}v_{\frac{1}{2}}}{\text{m.s.e. of } \delta^{2f-1}u_{\frac{1}{2}}} = \frac{1}{(4f-2,\ 2f-1)} \sum_{t=f}^{t=k} \left\{ \frac{\{2t-1, 2f-1\}}{\{2t-1,\ 2t-1\}} \frac{[\tfrac{1}{2}m,\ 2t)}{[\tfrac{1}{2}m, 2f)} \right\}^2 \frac{(4t-2,\ 2t-1)}{(m,\ 4t-1]}$$

$$= \frac{1}{(m,\ 4f-1]} \sum_{t=f}^{t=k} \frac{4t-1}{4f-1} \left\{ \frac{\{2t-1,\ 2f-1\}}{\{2f-1, 2f-1\}} \right\}^2$$

$$\frac{\{m^2-(2f)^2\}\ \{m^2-(2f+2)^2\} \ldots \{m^2-(2t-2)^2\}}{\{m^2-(2f+1)^2\}\ \{m^2-(2f+3)^2\} \ldots \{m^2-(2t-1)^2\}}; \quad (148)$$

$$\text{(E)}\quad \frac{\text{m.s.e. of } \mu\delta^{2f-2}v_{\frac{1}{2}}}{\text{m.s.e. of } \mu\delta^{2f-2}u_{\frac{1}{2}}} = \frac{4f-2}{(4f-4,\ 2f-2)} \sum_{t=f}^{t=k} \left\{ \frac{\{2t-2,\ 2f-2\}}{\{2t-2,\ 2t-2\}} \frac{(\tfrac{1}{2}m,\ 2t-1]}{(\tfrac{1}{2}m,\ 2f-1]} \right\}^2 \frac{(4t-4,\ 2t-2)}{(m,\ 4t-3]}$$

$$= \frac{4f-2}{(m,\ 4f-3]} \sum_{t=f}^{t=k} \frac{4t-3}{4f-3} \left\{ \frac{\{2t-2,\ 2f-2\}}{\{2f-2,\ 2f-2\}} \right\}^2$$

$$\frac{\{m^2-(2f)^2\}\ \{m^2-(2f+2)^2\} \ldots \{m^2-(2t-2)^2\}}{\{m^2-(2f-1)^2\}\ \{m^2-(2f+1)^2\} \ldots \{m^2-(2t-3)^2\}}. \quad (149)$$

23. *Smoothing.*—When we have a table containing a very large number of u's, a common method of procedure is to use a formula involving a limited number of terms and to apply it to successive sets of the u's for the purpose of obtaining a table to be substituted for the original table. Thus we might use a formula involving $2n+1$ terms, and apply it to $u_0, u_1, u_2, \ldots u_{2n}$ for finding a new value w_n, then to $u_1, u_2, u_3, \ldots u_{2n+1}$ for finding a new value w_{n+1}, and so on. These values having been obtained, a differenced table would be formed; but, as by hypothesis the true differences of order exceeding j are negligible, the table would only go up to differences of order j. There are two cases to be considered.

(i.) If our object is to obtain as accurate values as possible for the w's, consistently with our using only the specified number of u's for each, the most accurate values would be the v's given by the formulæ considered in this and the preceding papers. It should, however, be observed that the differences of the w's are not the same as the $\Delta^f v_0$, $\delta^{2f}v_0$, etc., occurring in those formulæ. Suppose, for instance, that we replace u_0 by its improved value v_0 obtained by means of the (B) formula involving $u_{-n}, u_{-n+1}, \ldots u_n$, and replace u_1 by the improved value v_1 obtained in a similar way. The resulting value of $v_1 - v_0$ will involve the $2n+2$ u's from u_{-n} to u_{n+1}; but it will not be the same thing as the improved value $v_{\frac{1}{2}}$ obtained by the (D) formula involving these u's, and its m.s.e. will therefore be greater than that of the latter.

(ii.) If our object is to obtain a smooth table of the w's as a whole, we could do this by obtaining as accurate values as possible for the differences of the w's of order j.

The formula which would have to be applied to the u's in order to obtain this result can be constructed without difficulty. The important thing to notice is that, if we alter the differences of the u's and then obtain the w's from the altered differences by summation, the resulting values must be such as can be legitimately substituted for the u's. Suppose, for instance, that $j = 2k+1$, and that we use $2n+1$ u's for each w. The formula for w will have to be of the form

$$w_0 = u_0 + c_{2k+2}\delta^{2k+2}u_0 + c_{2k+4}\delta^{2k+4}u_0 + \ldots + c_{2n}\delta^{2n}u_0\,;$$

and this will give

$$\delta^{2k+1}w_{\frac{1}{2}} = \delta^{2k+1}u_{\frac{1}{2}} + c_{2k+2}\delta^{4k+3}u_{\frac{1}{2}} + c_{2k+4}\delta^{4k+5}u_{\frac{1}{2}} + \ldots + c_{2n}\delta^{2n+2k+1}u_{\frac{1}{2}}.$$

The problem of determining the c's so that the m.s.e. of $\delta^{2k+1}w_{\frac{1}{2}}$ shall be a minimum is the same as that of determining the coefficients in the improved value of $\delta^{2k+1}u_{\frac{1}{2}}$ for $j = 4k+1$ or $4k+2$, m being $2n+2k+2$; and the solution of this problem is given in § 21. Thus, in terms of sums, (139) gives

$$\delta^{2k+1}w_{\frac{1}{2}} = \left[L\left\{-\sum_{g=1}^{g=2k+1}\lambda_{2k+1,\,2g-1}\mu\sigma^{2g}u_{\frac{1}{2}}\,;\ \sigma u_{t+\frac{1}{2}}\right\}\right]_{t=-(n+k+1)}^{t=n+k+1}.$$

The λ's having been found, we shall then have, by summation,

$$w_0 = \left[L\left\{-\sum_{g=1}^{g=2k+1}\lambda_{2k+1,\,2g-1}\mu\sigma^{2g+2k+1}u_0\,;\ \sigma^{2k+2}u_t\right\}\right]_{t=-(n+k+1)}^{t=n+k+1}.$$

The ratio of the m.s.e. of $\delta^{2k+1}w_{\frac{1}{2}}$ to that of $\delta^{2k+1}u_{\frac{1}{2}}$ is given by (148).

APPENDIX I.—THE CORRELATION-DETERMINANT.

1. The m.p.e. of A and B being denoted by $(A\,;\,B)$, let

$$\Theta \equiv \begin{vmatrix} (A\,;A) & (A\,;B) & (A\,;C) & \ldots \\ (B\,;A) & (B\,;B) & (B\,;C) & \ldots \\ (C\,;A) & (C\,;B) & (C\,;C) & \ldots \\ \vdots & \vdots & \vdots & \end{vmatrix}.$$

We call this the *correlation-determinant* of A, B, C, ...

2. The elements of this determinant may be regarded as obtained as follows. We first take a representative collection of N_A values of the error of A; N_A being usually indefinitely great. Then, for each of these values, we take a representative collection of N_B values of the error of B; the resulting N_A collections will all be alike if the errors of A and of B are independent, but not if they are correlated. This gives N_AN_B combinations of an error of A and an error of B. For each of these we take

a representative collection of N_C values of the error of C; and so on. Thus finally we shall have $N \equiv N_A N_B N_C \dots$ combinations of an error of A, an error of B, &c. Numbering these 1, 2, ... N, and denoting the errors of A, of B, of C, ... by a, b, c, ..., the combinations will be a_1, b_1, c_1, ..., a_2, b_2, c_2, ..., ... a_N, b_N, c_N, ...; and we shall have

$$(A\ ;\ A) = (a_1^2+a_2^2+a_3^2+\dots+a_N^2)/N,$$
$$(A\ ;\ B) = (a_1b_1+a_2b_2+a_3b_3+\dots+a_Nb_N)/N,$$
&c.

3. Substituting these values in Θ, we find that, if there are m of the quantities $A, B, C, \dots$,

$$N^m\Theta = \begin{vmatrix} a_1^2+a_2^2+a_3^2+\dots & a_1b_1+a_2b_2+a_3b_3+\dots & a_1c_1+a_2c_2+a_3c_3+\dots & \&c. \\ a_1b_1+a_2b_2+a_3b_3+\dots & b_1^2+b_2^2+b_3^2+\dots & b_1c_1+b_2c_2+b_3c_3+\dots & \&c. \\ a_1c_1+a_2c_2+a_3c_3+\dots & b_1c_1+b_2c_2+b_3c_3+\dots & c_1^2+c_2^2+c_3^2+\dots & \&c. \\ \vdots & \vdots & \vdots & \vdots \end{vmatrix}$$

$$= (a_1b_2c_3\dots)^2 + (a_1b_2c_4\dots)^2 + (a_1b_3c_4\dots)^2 + (a_2b_3c_4\dots)^2 + \dots,$$

where $(a_1b_2c_3\dots)$ denotes $\begin{vmatrix} a_1a_2a_3\dots \\ b_1b_2b_3\dots \\ c_1c_2c_3\dots \\ \vdots \end{vmatrix}$. Hence Θ is not $= 0$ unless each of the determinants $(abc\dots)$ is $= 0$. This would be the case, for instance, if A were a constant, so that every a would be 0, or if there were a linear relation connecting the errors of A, B, C, ...

4. Let Φ be the correlation-determinant of A, B, ... P, Q, ..., and Ψ that of A, B, ... P.

(*a*) Suppose that $\Psi = 0$. Then, by § 3 of this Appendix, each of the determinants $(ab \dots p)$, where a, b, ... p are the errors of A, B, ... P, is 0. But these are the minors of the q's in the determinants $(ab \dots pq)$; and therefore these latter determinants are 0. Proceeding in this way, we see that the determinants $(ab\dots pq\dots)$ are all 0; and therefore $\Phi = 0$.

(*b*) Hence, if Φ is not $= 0$, Ψ is not $= 0$.

Appendix II.—Frequency of Correlated Errors.

1. Let $u_0, u_1, u_2, \dots u_l$ and $y_0, y_1, y_2, \dots y_l$ be two conjugate sets. Denote the errors of the u's by $\theta_0, \theta_1, \theta_2, \dots \theta_l$; and let the resulting errors of the y's be $\phi_0, \phi_1, \phi_2, \dots \phi_l$.

Then, on the assumption of normal correlation of errors, the frequency of joint occurrence of these θ's is proportional to

$$\exp -\tfrac{1}{2}P,$$

where P is a homogeneous quadratic function of the θ's. We want to prove that

(i.) the ϕ's are the partial differential coefficients of $\frac{1}{2}P$ with regard to the θ's, and conversely;

(ii.) $P = \theta_0\phi_0+\theta_1\phi_1+\theta_2\phi_2+\ldots+\theta_l\phi_l$;

(iii.) $P = \psi_{0.0}\theta_0^2+2\psi_{0.1}\theta_0\theta_1+\psi_{1.1}\theta_1^2+\ldots+\psi_{l.l}\theta_l^2$, where $\psi_{f.g}$ is the m.p.e. of y_f and y_g; and similarly

(iv.) $P = \pi_{0.0}\phi_0^2+2\pi_{0.1}\phi_0\phi_1+\pi_{1.1}\phi_1^2+\ldots+\pi_{l.l}\phi_l^2$, where $\pi_{f.g}$ is the m.p.e of u_f and u_g.

2. Suppose that

$$P = a_{0.0}\theta_0^2+2a_{0.1}\theta_0\theta_1+a_{1.1}\theta_1^2+\ldots+a_{l.l}\theta_l^2;$$

and let us, without making any assumption of conjugacy, write ($f = 0, 1, 2, \ldots l$)

$$\begin{aligned} y_f &\equiv a_{f.0}u_0+a_{f.1}u_1+a_{f.2}u_2+\ldots+a_{f.l}u_l, \\ \phi_f &\equiv \text{error of } y_f \\ &= a_{f.0}\theta_0+a_{f.1}\theta_1+a_{f.2}\theta_2+\ldots+a_{f.l}\theta_l \\ &= \tfrac{1}{2}dP/d\theta_f. \end{aligned}$$

Then, writing the subscripts in the order f, 0, 1, 2, ... l,

$$\begin{aligned} P &= a_{f.f}\theta_f^2+2a_{f.0}\theta_f\theta_0+a_{0.0}\theta_0^2+\ldots+a_{l.l}\theta_l^2 \\ &= \phi_f^2/a_{f.f}+Q, \end{aligned}$$

where Q does not contain θ_f.

3. The mean value of $\phi_f\theta_g$ is N_g/D, where

$$N_g \equiv \iiint\ldots\int\phi_f\theta_g\exp-\tfrac{1}{2}P\,.\,d\theta_f\,d\theta_0\,d\theta_1\ldots d\theta_l,$$

$$D \equiv \iiint\ldots\int\exp-\tfrac{1}{2}P\,.\,d\theta_f\,d\theta_0\,d\theta_1\ldots d\theta_l,$$

the integration being in each case from $-\infty$ to ∞. If we write

$$\psi \equiv \phi_f/\sqrt{a_{f.f}},$$

then

$$N_g = \iiint\ldots\int\psi\theta_g\exp-\tfrac{1}{2}\psi^2\,.\,\exp-\tfrac{1}{2}Q\,.\,d\psi\,d\theta_0\,d\theta_1\ldots d\theta_l.$$

(a) First, suppose that g is not $=f$. Then, integrating with regard to ψ,

$$N_g = 0.$$

(*b*) Next, suppose that $g = f$. Then

$$N_f = 1/a_{f,f}\,.\iiint\ldots\int\phi_f(\phi_f - a_{f,0}\theta_0 - a_{f,1}\theta_1 - \ldots - a_{f,l}\theta_l)\exp{-\tfrac{1}{2}P}\,.\,d\theta_f\,d\theta_0\,d\theta_1\ldots d\theta_l$$

$$= 1/a_{f,f}\,.\iiint\ldots\int\phi_f^2\exp{-\tfrac{1}{2}P}\,.\,d\theta_f\,d\theta_0\,d\theta_1\ldots d\theta_l,$$

by (*a*). Hence

$$N_f = 1/\sqrt{a_{f,f}}\,.\iiint\ldots\int\psi^2\exp{-\tfrac{1}{2}\psi^2}\,.\exp{-\tfrac{1}{2}Q}\,.\,d\psi\,d\theta_0\,d\theta_1\ldots d\theta_l$$

$$= \sqrt{(2\pi/a_{f,f})}\,.\iint\ldots\int\exp{-\tfrac{1}{2}Q}\,.\,d\theta_0\,d\theta_1\ldots d\theta_l.$$

Also

$$D = 1/\sqrt{a_{f,f}}\,.\iiint\ldots\int\exp{-\tfrac{1}{2}\psi^2}\,.\exp{-\tfrac{1}{2}Q}\,.\,d\psi\,d\theta_0\,d\theta_1\ldots d\theta_l$$

$$= \sqrt{(2\pi/a_{f,f})}\,.\iint\exp{-\tfrac{1}{2}Q}\,.\,d\theta_0\,d\theta_1\ldots d\theta_l.$$

Hence

$$N_f/D = 1.$$

4. Hence the y's are related to the u's in such a way that

$$\text{m.p.e. of } y_f \text{ and } u_g = 0\ (g \neq f) \text{ or } 1\ (g = f);$$

and therefore the u's and the y's are conjugate sets; which proves (i.). It follows that

$$a_{f,g} = \text{m.p.e. of } y_f \text{ and } y_g.$$

This proves (iii.); and (iv.) is the similar result which we should have obtained by expressing P in terms of the ϕ's. Also

$$\begin{aligned} P &= a_{0,0}\theta_0^2 + 2a_{0,1}\theta_0\theta_1 + a_{1,1}\theta_1^2 + \ldots + a_{l,l}\theta_l^2 \\ &= \theta_0\,(a_{0,0}\theta_0 + a_{0,1}\theta_1 + a_{0,2}\theta_2 + \ldots + a_{0,l}\theta_l) \\ &\quad + \theta_1\,(a_{1,0}\theta_0 + a_{1,1}\theta_1 + a_{1,2}\theta_2 + \ldots + a_{1,l}\theta_l) \\ &\quad + \ldots \\ &\quad + \theta_l\,(a_{l,0}\theta_0 + a_{l,1}\theta_1 + a_{l,2}\theta_2 + \ldots + a_{l,l}\theta_l) \\ &= \theta_0\phi_0 + \theta_1\phi_1 + \theta_2\phi_2 + \ldots + \theta_l\phi_l; \end{aligned}$$

which proves (ii.).

APPENDIX III.—IMPROVED ADVANCING DIFFERENCES IN TERMS OF SUMS.

The expression for $\Delta^f v_0$ given by (136) differs from that given in (15) and § 5 of "Fitting," in that it involves $\Sigma'' u_0$, $\Sigma''^2 u_1$, $\Sigma''^3 u_2$, …, instead of $\Sigma'' u_0$, $\Sigma''^2 u_0$, $\Sigma''^3 u_0$, ….

The new expressions are more convenient for calculation and for tabulation, since the coefficients are rather smaller and are symmetrically placed about a diagonal. For $j = 2$, $m = 13$, for instance, the formulæ given by "Fitting," § 5 (ii.), are

$$1001v_0 = +693S_1-198S_2+22S_3,$$

$$1001\Delta v_0 = -231S_1+88S_2-11S_3,$$

$$1001\Delta^2 v_0 = +35S_1-15S_2+2S_3,$$

where $S_1 \equiv \Sigma'' u_0$, $S_2 \equiv \Sigma''^2 u_0$, $S_3 \equiv \Sigma''^3 u_0$. If we write $\Sigma_1 \equiv \Sigma'' u_0$, $\Sigma_2 \equiv \Sigma''^2 u_1$, $\Sigma_3 \equiv \Sigma''^3 u_2$, these become (by (136), or by writing $S_1 = \Sigma_1$, $S_2 = \Sigma_1+\Sigma_2$, $S_3 = \Sigma_1+2\Sigma_2+\Sigma_3$)

$$1001v_0 = +517\Sigma_1-154\Sigma_2+22\Sigma_3,$$

$$1001\Delta v_0 = -154\Sigma_1+66\Sigma_2-11\Sigma_3,$$

$$1001\Delta^2 v_0 = +22\Sigma_1-11\Sigma_2+2\Sigma_3.$$

The symmetry of the coefficients is due to the fact that

$$\text{co. } \Sigma''^{g+1} u_g \text{ in } \Delta^f v_0 = \lambda_{f,g} = \lambda_{g,f} = \text{co. } \Sigma''^{f+1} u_f \text{ in } \Delta^g v_0.$$

For any particular value of m there will be only $\frac{1}{2}(j+1)(j+2)$ coefficients to be tabulated, instead of $(j+1)^2$.

APPENDIX IV.—FORMULÆ IN TERMS OF u'S.

(i.) Formulæ for $\Delta^f v_0$, &c., in terms of the u's have already been given in (15a), (21), (22), (29), and (28) of "Fitting"; and the results in (141)–(145) of the present paper can be checked by comparing the different expressions for the coefficients of the u's. We should require to use the following identities :—

$$(r+h,\ h) = (r,\ 0)\ (h,\ 0)+(r,\ 1)\ (h,\ 1)+(r,\ 2)\ (h,\ 2)+\ \dots,$$

$$(r,\ 2h] = (0,\ 2h]+[r,\ 2)\ (0,\ 2h-2]+[r,\ 4)\ (0,\ 2h-4]+\ \dots,$$

$$[r,\ 2h-1) = (r,\ 1]\ [\pm\tfrac{1}{2},\ 2h-2)+(r,\ 3]\ [\pm\tfrac{1}{2},\ 2h-4)+(r,\ 5]\ [\pm\tfrac{1}{2},\ 2h-6)+\ \dots,$$

$$(r-\tfrac{1}{2},\ 2h-1] = [r-\tfrac{1}{2},\ 1)\ (0,\ 2h-2]+[r-\tfrac{1}{2},\ 3)\ (0,\ 2h-4]+[r-\tfrac{1}{2},\ 5)\ (0,\ 2h-6]+\ \dots,$$

$$[r-\tfrac{1}{2},\ 2h-2) = [\pm\tfrac{1}{2},\ 2h-2)+(r-\tfrac{1}{2},\ 2]\ [\pm\tfrac{1}{2},\ 2h-4)+(r-\tfrac{1}{2},\ 4]\ [\pm\tfrac{1}{2},\ 2h-6)+\ \dots.$$

(ii.) Taking, for instance, the formula for $\delta^{2f}v_0$ when $m = 2n+1$, (21) of "Fitting" gives (replacing t by f)

$$(p_r)_{2k} = (-)^f \tfrac{1}{2}m \frac{\{2k+1,\ 2f+1\}}{[\frac{1}{2}m,\ 2f+1)} \sum_{h=0}^{h=k} (-)^h \frac{(f+1)(h+1)}{f+h+\frac{1}{2}} \frac{\{2k+1,\ 2h+1\}\ (r,\ 2h]}{(\frac{1}{2}m,\ 2h+1]}.$$

Hence we find that

$$(p_r)_{2k}-(p_r)_{2k-2} = (-)^f\tfrac{1}{2}m\,(2k+\tfrac{1}{2})\frac{\{2k,\,2f\}}{[\tfrac{1}{2}m,\,2f+1)}\sum_{h=0}^{h=k}(-)^h\frac{\{2k,\,2h\}\,(r,\,2h]}{(\tfrac{1}{2}m,\,2h+1]},$$

$$(p_r)_{2k} = (-)^f\tfrac{1}{2}m\sum_{t=f}^{t=k}(2t+\tfrac{1}{2})\frac{\{2t,\,2f\}}{[\tfrac{1}{2}m,\,2f+1)}\sum_{h=0}^{h=t}(-)^h\frac{\{2t,\,2h\}\,(r,\,2h]}{(\tfrac{1}{2}m,\,2h+1]};$$

and therefore

$$(\delta^{2g}p_0)_{2k} = (-)^{f+g}\tfrac{1}{2}m\sum_{t=f}^{t=k}(2t+\tfrac{1}{2})\frac{\{2t,\,2f\}}{[\tfrac{1}{2}m,\,2f+1)}\frac{\{2t,\,2g\}}{(\tfrac{1}{2}m,\,2g+1]}\,\Theta,$$

where

$$\begin{aligned}\Theta &= F\{-t+g,\ g+t+\tfrac{1}{2},\ \tfrac{1}{2};\ 1,\ -\tfrac{1}{2}m+g+1,\ \tfrac{1}{2}m+g+1\}\\ &= \frac{(\tfrac{1}{2}m,\,2g+1]\,[\tfrac{1}{2}m,\,2t+1)}{(\tfrac{1}{2}m,\,2t+1]\,[\tfrac{1}{2}m,\,2g+1)}.\end{aligned}$$

This expression for $(\delta^{2g}p_0)_{2k}$ will be found to be equal to $\lambda_{2f,2g}$ as given by (132) of the present paper, so that the formula in " Fitting " agrees with (142).

VIII. *Tidal Friction in Shallow Seas.*

By HAROLD JEFFREYS, *M.A.*, *D.Sc.*, *Fellow of St. John's College, Cambridge.*

Communicated by Sir NAPIER SHAW, *F.R.S.*

Received April 7,—Read June 24, 1920.

THE astronomical importance of the dissipation of energy that goes on in shallow seas has been shown by G. I. TAYLOR'S recent estimate* of the amount in the Irish Sea, which is enough to account for about one-fiftieth of the secular acceleration of the moon. It also produces a considerable effect on the tides themselves, and there are probably many places where it must be taken into account before any satisfactory theory of the local tides, or even their empirical prediction, can be achieved. It is indeed very well known that there are bays and straits where the height of the tides, or the speed of the currents, or both, are greater than in the Irish Sea, and a careful examination of such places, with a view to finding the dissipation in them, is needed. There are other places where the dissipation for an equal area is less than in the Irish Sea, but which may actually contribute much more altogether on account of their greater size. The object of this paper is to discuss what regions are capable of producing notable parts of the secular acceleration; to estimate as accurately as possible from the data available the dissipation in these; and to compare this with that calculated from the secular acceleration, so as to find out whether it is necessary to assume the existence of any other important cause to account for the latter.

The horizontal force of the skin friction of water over the sea bottom is $0{\cdot}002\rho V^2$ dynes per square centimetre, where ρ is measured in grammes per cubic centimetre and V in centimetres per second. The difficulty of the problem is in the estimation of V. The available observations of the velocities of tidal currents are given in the Admiralty Sailing Directions; but they are never uniformly distributed, and are usually confined to the neighbourhood of the coasts, and they must be supplemented by theory before the velocities remote from the coast can be found. A few theoretical considerations that have been found useful in this process will now be mentioned.

Take first the case of a bay or strait long in comparison with its width, and consider a wave entering it whose period is much longer than the time needed for a

---

* 'Phil. Trans.,' A, vol. 220, pp. 1–33, 1919.

[Published November 5, 1920.

wave of tidal type to cross. This time is known to be independent of the period of the disturbance, depending only on the width and depth. Such a wave is reflected from one side to the other again and again before it reaches the other end. It is also known that the transverse velocity at the sides is zero, since water cannot cross a rigid boundary. Thus if we compare two points on opposite sides of the channel, we know that the times of arrival of the wave at them differ by a small fraction of a period; and since the transverse velocity at both is zero, it cannot be great at any intermediate point, for that would contradict the hypothesis that the wave-length is much greater than the width of the channel. The transverse velocity may accordingly be neglected in problems of this class.

If the period of the entering wave is of the same order of magnitude as the time needed to cross the channel, we can no longer infer that the transverse velocity is much less than the longitudinal one. This case seldom or never arises. The velocity of a tidal wave is $(gD)^{\frac{1}{2}}$ where g is the intensity of gravity and D is the depth; and if the water was only 20 fathoms deep a tidal wave would in 12 hours travel 700 km., which is far greater than the width of almost any channel whose length is much greater than its width. Where the width is greater, the depth also is always greater, so that the above argument always holds in long channels of whatever size.

If now x be the distance of a point from the entrance to the channel, y the distance from the side, u the longitudinal velocity of a particle there, and η the height of the free surface above its undisturbed position, the equations of motion of the particle are

$$\frac{du}{dt} = -\frac{g\partial\eta}{\partial x} - \text{a term due to friction}$$

$$2\omega u = -g\frac{\partial\eta}{\partial y}$$

where ω is the component of the earth's angular velocity of rotation about the vertical at the point. The equation of continuity is

$$\frac{\partial}{\partial x}(\mathrm{D}u) = -\frac{\partial\eta}{\partial t}.$$

From this and the second equation of motion we deduce at once

$$\frac{\partial\eta}{\partial x} = \frac{\partial\eta_0}{\partial x} + \frac{2\omega}{g}\int\frac{1}{\mathrm{D}}\left(\frac{\partial\eta}{\partial t} + u\frac{\partial\mathrm{D}}{\partial x}\right)dy$$

where η_0 is the value of η at the side. In this we see from the conditions that the channel is narrow and the depth slowly varying along it that the first term is much greater than the others. Accordingly $\frac{\partial\eta}{\partial x}$ is the same for all particles in the same cross-section of the channel, and the first equation of motion then shows that the same is true of u if the friction is small. Thus in such cases the velocity is the same at all points in the same cross-section, so that observations made at the side will be correct for points in the middle. If friction is great this result must be modified, since for the same velocity the frictional force is independent of the depth of the water, whereas the mass affected is proportional to the depth. Thus friction has more influence in reducing the velocities in shallow water than in deep water. Hence when the channel is shallow at the sides and deep in the middle the velocity will be least at the sides. The ratio of the frictional term to the first term is $0{\cdot}002\, u/2D'\Omega$, where D' is the depth at the shallow part. When this is 10 fathoms and the velocity 1 knot the fraction is 0·25, so that this effect is then appreciable; and when the depths are less and the velocities greater, the influence of friction may increase to such an extent as to dominate the whole character of the motion. When this occurs, the velocity will always be in the direction of decreasing pressure, and inertia may be neglected.

The last result may appear to contradict the general principle that "still waters run deep." There is no real contradiction, however, for the problems referred to are different. The above argument deals with the differences between the velocities at places in the same cross-section of the channel, whereas the proverb concerns rivers whose depth varies *along* them, and in such cases the motion is naturally slowest where the depth is greatest, since the amount of water crossing any section in a given time must be the same. It also has an important and well-known application in bays of varying depth and width, such as the Bay of Fundy. If, for instance, the bay is very long, and these quantities change only by small fractions of themselves per wave-length, it can be shown* that the height of the wave at any point is proportional to $b^{-\frac{1}{2}} D^{-\frac{1}{4}}$, and the velocity of the water to $b^{-\frac{1}{2}} D^{-\frac{3}{4}}$, where b is the width at the surface. The rate of dissipation of energy across any section is proportional to bu^3 or to $b^{-\frac{1}{2}} D^{-\frac{9}{4}}$. It therefore increases slowly as the channel becomes narrower and much more quickly as it becomes shallower. When the depth and width vary much within a wave-length these results cease to be useful approximations, but the tendency for the height of the tide and the velocity of the tidal current to increase as the channel becomes narrower and shallower remains. Thus in such places we often find very high tides and strong tidal currents. Apparently, however, their limited area prevents the dissipation in them from being as great as that in larger places with less violent currents (at least, if the Bay of Fundy may be regarded as typical of them).

The widths of most actual bays are, however, comparable with their lengths, and in these it is generally a matter of some difficulty to settle whether we can treat the recorded currents as a fair sample of the whole. The amplitude of the tide in mid-ocean is only about a foot, but in the shallow water around the coasts it is magnified to several feet, and the tidal currents are increased correspondingly. Where the shore is fairly open and regular in outline, like most of the coast of Africa, it is not possible to find the dissipation along it, for there are no data to show how far out the currents extend. In partly enclosed regions, however, it is frequently possible to interpolate between the records made on opposite sides. A serious difficulty may arise if the depth of the sea is very different at different points within it, for this may destroy the possibility of interpolation, and therefore we must always examine the soundings for any great variation. Ordinarily we should not expect much variation in the velocities, for such places are intermediate in character between narrow channels and the open shore, and therefore the currents in them may be expected to show some increase in shallow water, but not so much as would be caused by a proportional decrease in depth along a narrow bay or in approaching the open shore. In shallow water also friction may, and often does, neutralize the magnification that would occur in its absence.

---

* LAMB, 'Hydrodynamics,' p. 258.

One other fact may be noted. In shallow bays the difficulties of navigation may be great, and navigators avoid them if possible by choosing a harbour near the entrance. Thus observations of currents are most numerous about the entrances, and often at the very places where the currents, and consequently the dissipation, are greatest there are insufficient observations to give a satisfactory estimate.

Great care must be taken in dealing with observations among islands, straits, and shoals. When the passage of a tidal current is obstructed by a shallow of small horizontal extent, part of it goes round the shoal and part over the top. The influence of this on the main current is of course small, but on the top of the shoal and in its immediate neighbourhood the velocities may be much increased, for much the same reason as accounts for the greater speed of a river where it is shallow. On the other hand the increase in the influence of friction may greatly reduce the currents, and shoals often afford in this way an important shelter from tidal currents to the deeper water behind them. This is particularly noticeable at some points on the Korean side of the Yellow Sea. Thus observations of currents taken at lightships and buoys over shoals whose dimensions are all much smaller than those of the main bay or channel must be regarded as giving no reliable estimate of the main current. Small islands also require examination before the records obtained are accepted. If one is surrounded by a shoal they are of course untrustworthy; but if deep soundings are found within a few miles of it, they will probably give a very good idea of the main current, which will, especially in a wide channel, be fully as useful in our investigation as the results of observations at the sides. Straits are in a different position. When the tides in two seas or even oceans are in widely different phases, a large head may be produced between the two ends of a strait connecting them, so that a swift tidal current will flow along the strait. In no circumstances, however, can this give any indication of the currents in the seas, for it is produced by the tide heights, and not directly by the currents. Such currents may attain very great velocities, as in the

Magellan Strait and Smith Sound, and when the area of the strait is not insignificant the dissipation may be an important part of that in the seas as a whole.

EUROPEAN SEAS.

1. *The Irish Sea.*

This sea has been discussed in detail by TAYLOR. The rate of dissipation is found to be 1040 ergs per square centimetre per second, or $4{\cdot}1 \times 10^{17}$ ergs per second in all, on an average at spring tides. This result is based on the law that the rate is proportional to the cube of the velocity. The Irish Sea is remarkable in that the maximum current occurs nearly at high water, whereas in ordinary places the water is nearly slack then; though other examples will be given later in this paper. This affords the most favourable conditions for an accurate estimate of the rate at which energy enters the sea; to this TAYLOR added the rate at which the moon's attraction does work on the sea, and from the fact that all this energy must be dissipated in the course of a period (for if it were not, there would be a continual increase of energy in the Irish Sea) he found that the mean rate of dissipation at spring tides was $6{\cdot}0 \times 10^{17}$ ergs per second. This estimate is probably more accurate than the other, as the data involved are obtained from observations in St. George's Channel, supplemented by an accurate theory; but the former is based on an average of the velocities in the Irish Sea itself, which are more difficult to determine.

2. *The English Channel.*

On an average the tidal currents in the English Channel at springs reach about 2·5 knots. The speed is greatest towards the Straits of Dover and least at the entrance to the Channel, and enough data are available in the Admiralty publication, 'The Tides and Tidal Streams of the British Isles' to give a very accurate estimate of the total dissipation if this were required in the present problem; but as the errors introduced by using only a rough approximation in this case are far less than those involved in the best data referring to regions with far larger dissipations, accuracy is not here required.

The rate of dissipation is $0{\cdot}002\rho V^3$ ergs per square centimetre per second. Here ρ is practically 1; and one knot is 51·5 cm. per second. Thus the dissipation per square centimetre for a velocity of one knot is 274 ergs per second, and that per square kilometre is $2{\cdot}74 \times 10^{12}$ ergs per second. The area of the English Channel is about 60,000 sq. km., so that the dissipation when the currents are flowing fastest is $2{\cdot}74 \times 10^{12} \times 6 \times 10^4 \times 2{\cdot}5^3$, or $2{\cdot}5 \times 10^{18}$ ergs per second. This is of course a maximum, while the value obtained for the Irish Sea is the mean over a period; the average rates of dissipation in the two places are perhaps not very different.

3. *The North Sea.*

A satisfactory estimate of the dissipation in the North Sea is practically impossible. Velocities up to over 3 knots are recorded here and there, but all the observations are in the coastal region, which is very much complicated by shoals. The maximum in the outermost part of the Moray Firth is about 1·1 knot, and this is probably fairly typical of the whole of the North Sea. Taking the area to be 5×10^5 sq. km. and adopting the above value of the velocity, we see that the maximum dissipation is of the order of $1{\cdot}8 \times 10^{18}$ ergs per second.

4. *Other European Waters.*

In the Mediterranean there is probably little or no dissipation of tidal energy, for the Atlantic tidal wave can only enter through the very narrow Straits of Gibraltar, and partly for this reason and partly on account of the great length and considerable depth of the sea there is very little tidal movement in it. The same argument applies to the Baltic, for the entrance through the Kattegat is largely blocked up by the Danish islands, so that little water can enter to produce a tide. The Bay of Biscay is mostly too deep to have any important current, while the White Sea is too small and landlocked to give as much dissipation as the Irish Sea.

The average dissipation in a period is $4/3\pi$ of the maximum. If we find the maximum for the Irish Sea on this basis, we obtain for the total dissipation in European waters when the spring tide currents are flowing strongest about $6{\cdot}0 \times 10^{18}$ ergs per second; the average at spring tides is $2{\cdot}4 \times 10^{18}$ ergs per second.

ASIATIC SEAS.

It has already been pointed out that the tidal currents in mid-ocean are insufficient to give any important dissipation.* Accordingly we need consider only those places where the currents are very much magnified by great decreases in depth. On referring to a physical map of Asia it is at once seen that the places around the coast where the depth is less than 100 fathoms are the Straits of Malacca, the South China Sea (with the Java Sea), the Gulfs of Siam and Tongking, the Yellow Sea, the Persian Gulf, and parts of the Seas of Japan and Okhotsk and the Bering Sea. These regions will be dealt with separately. The Persian Gulf may be omitted at once, as its narrow entrance prevents the tide from being great.

1. *The South China Sea.*

This sea is in the form of a letter T. The middle stroke points north-east and lies between Annam and Southern China on the one side, and Borneo and the Philippines on the other. The two side-pieces are the Gulf of Siam and the Java Sea. The data for it are obtained from the 'China Sea Pilot,' volumes 3 and 4, except the depths, which are taken from the Admiralty Charts. The tides are affected by a large diurnal inequality due to the inclination of the Equator to the ecliptic. When the moon is north of the Equator it tends to raise two tidal protuberances in the ocean, one exactly below it and the other exactly opposite to it. Owing to the earth's rotation each of these moves round the earth once a day, keeping the same distance north or south of the Equator. Thus if a place is not on the Equator, they pass at different distances from it, so that the two tides in the lunar day are unequal in height. The variation in the level of the water thus caused can be described as a semi-diurnal change, on which a diurnal change is superposed.

---

* 'Philosophical Magazine,' May, 1920, vol. 39, pp. 578–586.

Now the rate of travel of a tidal wave is practically independent of its period, but if the depth and the form of the coast are such that the waves starting from the north and south sides of the Equator take different times to reach the place of observation, their combined effect may be remarkable. In particular, if the wave from the south arrives a quarter of a lunar day after or before the other, the semi-diurnal part of the one wave will correspond to high water while that of the other corresponds to low water, and if the amplitudes are equal the two will neutralize each other. In other words, there will be a node of the semi-diurnal tide. The diurnal parts, however, will not neutralize each other, their phases being only a quarter of a period apart. Thus at such a place there will be a diurnal tide and no semi-diurnal tide. Several places are known where there is only one high water in each lunar day; among them are parts of the South China Sea, the Gulf of Carpentaria, and Bering Sea. The tides in these require special discussion before they can be considered in the present problem, because the diurnal tide depends essentially on the inclination of the Equator to the plane of the moon's orbit and would not exist if this were zero. The dissipation of energy in it must therefore arise from the motion of the moon in declination and not in right ascension, and will affect mainly the inclinations of the Equator and the moon's orbit to the ecliptic, while producing little effect on the earth's rotation and the motion of the moon in longitude. In discussing the secular acceleration of the moon it can therefore be ignored. If observations of the diurnal tide in the places where there are two high waters in the day were more numerous it might be possible to determine the dissipation in it, and from it the secular changes in the inclinations, but at present this is impossible.

In the Java Sea, between Borneo and Java, the tide is mainly diurnal; in fact according to the 'Eastern Archipelago Pilot,' part 3, the semi-diurnal tide is not appreciable on the north coast of Java till east of Surabaya, which is itself almost at the eastern end. On the south coast of Borneo the observations are not so numerous, but it seems clear that there also the tide is mainly or entirely diurnal. The tidal currents are described as weak. We can accordingly neglect the dissipation in the Java Sea. At its western end this sea is connected to the South China Sea by two rather wide straits, Carimata Strait and Gaspar Strait. In both of these the tide is mainly diurnal; in the latter, in fact, it is entirely so. At Pontianak, near the north-westerly point of Borneo, the tide is still diurnal; thus the part of the China Sea south of Pontianak and Singapore probably contributes little to the secular acceleration of the moon.

Apparently the main tidal stream from the China Sea strikes the Malay Peninsula somewhere near Cape Patani and spreads out from there; for on the coast north of this point the flood stream sets to the north, while south of it it sets to the south. The tide in this region is definitely semi-diurnal, though the heights of the two daily high waters may be unequal. The depths in the western part of the sea and in the Gulf of Siam are mostly about 30 fathoms, but there are many shoals around the coast where the depth is only a few fathoms, and it is therefore necessary to be very critical of the sites of observations of currents. The best results seem to be given at small islands with rapidly shelving sides, for the currents there are modified little by the form of the bottom and can be regarded as fairly typical of the general currents in the neighbourhood.

At the Anamba Islands (see map, fig. 1) the semi-diurnal tide appears to be usually much less than the diurnal one. The 'China Sea Pilot,' vol. 3, states that for a few days in each month, when the moon is near the Equator, there are two high tides in the day. It is easily seen that for the two tides to occur in the day the amplitude of the semi-diurnal term must be $4/3\pi$ of a quarter of that of the diurnal term. It would not, however, vary much with the moon's declination, whereas that of the diurnal term vanishes when the moon is on the Equator; and the above fact shows that the semi-diurnal tide only attains this fraction of the diurnal tide when the latter is at its least. The true semi-diurnal tide at the Anamba Islands must therefore be insignificant. The same is evidently true of the currents, for the tidal streams take a day to run backwards and forwards.

In the Gulf of Siam also the tides are mainly diurnal. The oscillation in this is a forced one due to that in the South China Sea, and as the latter is diurnal so is that in the Gulf of Siam. Actually the only place where the semi-diurnal tide is considerable is Bangkok Harbour, at the head of the gulf. This tide seems to increase in relative importance towards Bangkok, for at Kamput on the eastern side and places in about the same latitude on the western side the tides are said to be very irregular, indicating the presence of some complicating influence. On the whole, therefore, it seems that the dissipation in the Gulf of Siam will not be underestimated if we assume that the semi-diurnal current reaches a maximum of one knot, this being one-third of the diurnal tidal current observed off Cape Patani, at all places north of the parallel of 11° N. The area of this region is about 70,000 sq. km., giving a maximum dissipation of 2×10^{17} ergs per second.

An estimate may be made of the dissipation of the energy of the diurnal tide in the same regions. At Pontianak there is a diurnal current of two knots when running strongest; and a similar velocity is recorded at the Burong Islands some distance to the north. As the depth of the passage between Borneo and Sumatra is fairly uniform, from 10 to 20 fathoms, these measures are probably typical of the whole. Quantitative estimates of the currents in the Gulf of Siam are few, and a reasonable guess at them would be difficult. The depth is mostly about 30 fathoms, but there are a deeper area in the middle and many shoals about the margins. As the current at the mouth of the gulf is across it, the tide in the gulf can arise only from the reflection of this by the Malay coast, so that the general set of the current is across the gulf, and considerable magnification in the gulf is unlikely. Our estimate of the dissipation will probably be of the correct order of magnitude, if we suppose that everywhere west of a line joining Cape Datu to Cambodia Point the maximum current is two knots. The area of this region is 8×10^5 sq. km.; thus the maximum dissipation is $1{\cdot}7 \times 10^{19}$ ergs per second. The velocity is proportional to the sine of the hour angle of the moon increased by a constant; and as the dissipation is proportional to the cube of the velocity, the average dissipation is obtained by

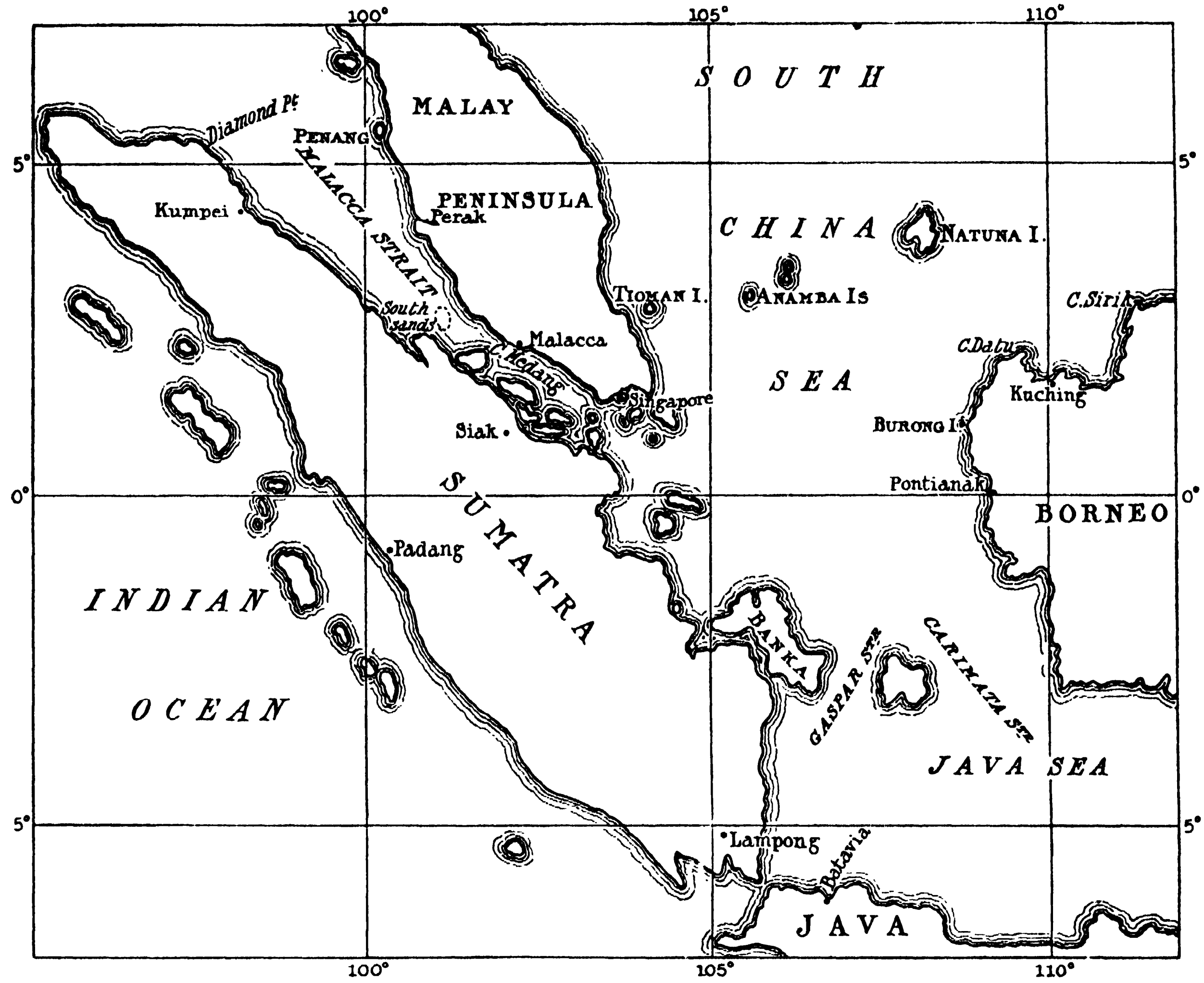


Fig. 1.

multiplying the maximum by the average arithmetical value of $\sin^3 \theta$ taken over a period, which is $4/3\pi$. It is not clear what declination of the moon the recorded currents refer to: if they refer to the maximum, the average for the month will be found by multiplying again by $4/3\pi$. Thus the average dissipation of energy in the diurnal tide in the western part of the China Sea is of the order of 3×10^{18} ergs per second.

We next proceed to examine the dissipation in the main part of the South China Sea, between the north-west coast of Borneo and the mainland. The currents are now semi-diurnal. Near Cape Sirik the flood runs for four hours and the ebb for eight hours, so that a considerable diurnal component exists, but not sufficient to preponderate over the semi-diurnal motion. The velocities here are from two to three knots. At Bruni the observations are very much interfered with by shoals and narrows, but in the offing the currents seem to be about two to three knots. About Tega, however, among shoals the velocities recorded are only 1·3 and 0·8 knots. The contours of the sea floor run roughly parallel to the coast, so that these currents may persist for some distance out to sea. Off the north end of Borneo the sea rapidly deepens, and in accordance with this there is scarcely any tidal current at Ulugan Bay, in Palawan. On the Asiatic side the tide is diurnal at Camran and Tourane, but the currents are weak.

A few shoals and islands in the middle of the sea have been made the localities of observations. At Rifleman Bank and Spratly Island there is only one tide in the day, and at the neighbouring island of Amboyna it is said that near neaps the stream reaches 1·4 knots. It is therefore clear that the semi-diurnal current of the Borneo coast does not extend half way across the sea, and its true extent is very doubtful. In the Gulf of Tongking also the currents appear to be diurnal. Thus in the whole of the South China Sea and its extensions there seems to be little semi-diurnal tide and little contribution to the secular acceleration of the moon, though there is a dissipation of the energy of the diurnal tide that may have a notable secular effect on the obliquity of the ecliptic.

2. *The Yellow Sea.*

This is a gulf about the size of Ireland, lying between Korea and the coast of China, and extending about as far south as the mouth of the Yang-tse-Kiang. It becomes very narrow where the Shan-tung peninsula projects into it, and north of this it forms the Gulfs of Pe-Chili and Liau-tung. Most of it is shallow, the depths in the main part of it being mostly about 30 fathoms, and those in the northern part about 15 fathoms. Around the shore the water is shallower, and in many places there are crowds of shoals. The data are obtained from the 'China Sea Pilot,' vol. 5.

The tidal phenomena are extremely complex. At the south end of the peninsula of Korea high water (full and change) is at about 11h., Greenwich time. As we advance up the Korean coast it occurs later and later, being practically 12 hours later at Port Arthur. On the opposite side of the Strait of Pe-Chili it is slightly earlier than at Port Arthur; then on the way round the Shan-tung peninsula and out of the sea again the tide again becomes steadily later. It therefore looks as if the tide enters up the coast of Korea, gradually passes up the sea, losing energy all the way, and a reflected wave from the Pe-Chili strait emerges down the Chinese coast. The tides and the currents on the Korean side are noticeably stronger than those on the Chinese side, and it does not seem likely that this is due wholly, or even largely, to the shoals on the former coast, for at islands in deep water, such as the Mackau, Myangoru, and Bate groups, velocities of 3 to 5 knots are recorded, while in shallower water the velocities are not usually greater than these, though local strong streams exist. Apparently the effects of friction are great enough to counterbalance those of the diminution in depth. They are also seen in another respect. Among the islands of the Korean archipelago the tidal stream sets west from four hours before till two hours after high water, whereas in most places elsewhere there is little or no current at high water. Thus work is continually being done on the sea, and the energy entering is dissipated in it. An effect of the approximate agreement in phase between the tide and the velocity may be utilized to give an estimate of the currents and tide height in the entrance to the sea, far from the nearest land, and hence of the amount of energy entering. In any motion in a channel where, on

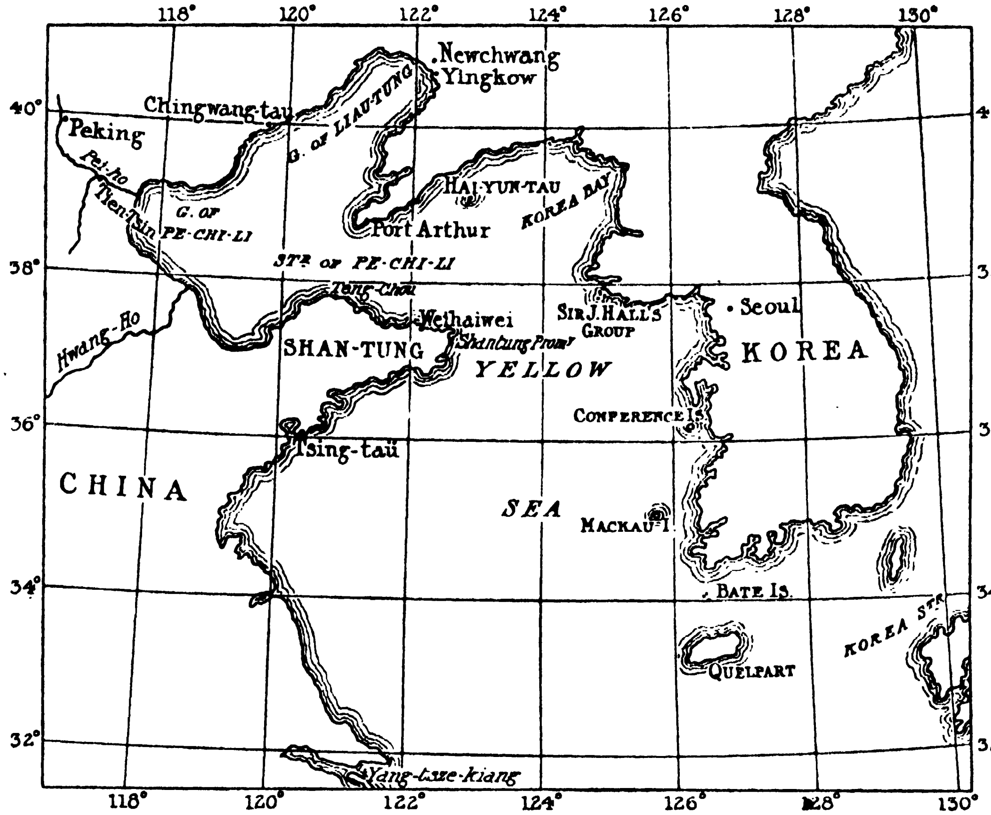


Fig. 2.

account of its narrowness or for some other reason, there is little transverse motion, the velocity along the channel is related to the pressure gradient across it according to the equation

$$2\omega\rho u = -\partial p/\partial y,$$

where ρ is the density of the water,

ω is the component of the earth's angular velocity of rotation about the vertical at the point considered,

p is the pressure, and

y is the distance measured across the channel.

In the Yellow Sea the entrance is not narrow, but there seems reason to believe that the velocity across it is small, which is all that is required for the truth of the above equation. If now g denote the intensity of gravity, and η the elevation of the surface of the water above its mean position, then at any fixed point, at whatever depth, the variation of p is equal to that of $g\rho\eta$. Again, if Ω be the earth's angular velocity, and λ the latitude of the place,

$$\omega = \Omega \sin \lambda,$$

and we have on putting $\lambda = 35°$; $\Omega = 7{\cdot}3\times 10^{-5}/1$ sec.; $g = 981$ cm./sec.^{2};

$$u = -1{\cdot}17\times 10^{7}\frac{\partial\eta}{\partial y}$$

where C.G.S. units must now be used.

On the coast of Korea the tide has an amplitude of about 10 feet, or 300 cm. The velocity of the inward current is about 4 knots, or 200 cm./sec. Now suppose if possible that the current remained constant right into the middle of the entrance; then the above formula shows that at a distance of 176 km. from the side there would be little vertical movement of the surface, and further away still a huge tide with an amplitude of some 30 feet would exist. It is not reasonable that the tide in the middle should be greater than that at the side, though it may easily be smaller. The alternative hypothesis is therefore that the current decreases as we approach the middle, and is very small over most of the sea. This will be adopted in the forthcoming discussion. We shall suppose that the current at distance y from the coast is in the same phase as that at the coast, and is a linear function of y. Then put

$$u = (200-ky)\cos\gamma t.$$

At the shore η is equal to $300\cos(\gamma t-\alpha)$, where α is the difference in phase between the tide height and the current strength. For the semi-durnal tide it is twice the angle moved through by the moon relatively to the earth in one hour, or 29 degrees. In general

$$\eta = 300\cos(\gamma t-\alpha)-8{\cdot}6\times 10^{-8}\cos\gamma t\,(200y-\tfrac{1}{2}ky^{2}).$$

The amplitude of η reaches a minimum where the coefficient of $\cos \gamma t$ vanishes. We shall have the least possible tide in the middle, thus satisfying our earlier assumptions, if we assume that this minimum is reached at the point where the velocity vanishes. In this way we shall underestimate the dissipation, but not by any great amount. The value of k that makes this coincidence possible is $6{\cdot}6 \times 10^{-6}/1$ cm., so that the current becomes zero 300 km. from the shore, practically in the middle of the entrance.

TAYLOR shows (*loc. cit.*, equation 15) that the average amount of energy crossing any line is the average over a period of $g\rho\int \mathrm{D}\eta u \, dy$ in my notation, where D is the depth of the water, in this case about 30 fathoms over most of the region in which the velocity is greatest, with a range in all of from 20 to 45 fathoms. We find easily

$$\eta = 147 \sin \gamma t - 4{\cdot}3 \times 10^{-8} \, k \, (y - 200/k)^2 \cos \gamma t,$$

and the average flux of energy across a parallel of latitude is found to be $1{\cdot}06 \times 10^{18}$ ergs/sec.

The northward flux on the Chinese side is more difficult to determine, as the direction of the currents is variable. Two phenomena are intermingled here. The issuing tide from the Yellow Sea comes down this coast, but there is also a definite tidal wave that travels into and out of the large bend in the coast whose extremities are Shan-tung promontory and Shanghai. This is shown by the fact that along the northern part of this bend, on which Tsing-tao stands, the current flows northwards while the tide is ebbing both along this coast and in the northern part of the Yellow Sea. Thus in this bay the main tide is the local tide of the bay itself and not the general tide of the Yellow Sea. The currents produced by these tides are rotary, probably an effect of the earth's rotation; and it seems that the northward component of the velocity is small. Further, it is probably nearly in a phase at right angles to the tide, as great divergences from this relation can be produced only by great dissipations and accordingly by great velocities in the vicinity. Hence for both reasons we infer that the northward flux of energy along the Chinese coast is small in comparison with that on the Korean side.

We also need to know the work done on the water by the moon. If η' is the height of the equilibrium tide, the work done by the moon in a period is $\int\int g\eta' \frac{d\eta}{dt} \, dt \, d\mathrm{S}$, where $d\mathrm{S}$ is the element of horizontal surface and the integrals are to be taken, in the one case over a period, and in the other over the area considered. If the phase of η is β in advance of that of η', and the amplitudes be h and h', the average rate of doing work is $-\frac{1}{2} g \int hh' \gamma \sin \beta \, d\mathrm{S}$. In the present case h' naturally varies little; h decreases fairly steadily as we travel from the entrance to Port Arthur, where it has about half of its value at the entrance. On the other hand β varies a great deal. The longitude being about 120° E., the moon crosses the meridian at full and change at 3h. 43m. It is high water in Shoan harbour, at the southern end of Korea, at 10h. 33m., so that β is here $-199°$, or more conveniently $+161°$. The tide lags more and more on the way up the sea, and at Port Arthur it has lost practically a whole period, the high water at full and change occurring there at 11h. 7m., making $\beta = -216°$. Positive elements will be added to the integral by the places where β is from 161° to 0°, negative elements where β is from 0° to $-180°$, and positive again where it is from $-180°$ to $-216°$. These values are however weighted according to the values of h and according to the extent of the areas for which they are correct. Now β is zero near the Conference Islands, only about 0·4 of the way to the narrowest part, so that on this account the area for which it is positive would appear to be less than that for which it is negative. The sea becomes much narrower farther north, however, which must reduce the ratio of the weights somewhat, and the tides are only about two-thirds of the height. Thus it seems that the weights to be attached to the positive and negative values of $\sin\beta$ are nearly equal, and its average over the sea is unlikely to be more than 0·2. The area of the sea as far as Port Arthur is about 300,000 sq. km.; taking the average of h as 240 cm., and that of h' as 20 cm., we find the energy imparted by the moon to be not greater in absolute magnitude than 2×10^{17} ergs per second on an average. That entering up the Korean coast is far greater.

In the discussion of the work done by the moon the Gulfs of Pe-Chili and Liau-tung have been ignored. There are two reasons for this: their united area is about a third of that of the main part of the sea, and the tides recorded at the sides are also about a third of those on the Korean coast. It seems to me, however, that these two gulfs afford an example of a special type of tidal problem different from any previously discussed. For, let us suppose if possible that the recorded amplitudes, of the order of 90 cm., were typical of the whole area. The average depth is about 2000 cm., and a tidal wave in water of such a depth would give rise to a current of maximum velocity about $h\,(g/\mathrm{D})^{\frac{1}{2}}$, which in this case is 65 cm./sec. The corresponding dissipation would be 550 ergs per square centimetre per second. On the other hand the average energy present is about $\frac{1}{2}g\rho h^2$, or 4×10^6 ergs per square centimetre. Thus if the above assumption were correct the whole energy of the tide would be dissipated in about 7000 seconds, or two hours. This is absurd, and we must suppose, in order to avoid the result that energy is dissipated faster than it enters the region, that the tides in the greater part of the gulfs are much less than the recorded ones. Their height may be only a few inches; while the recorded heights are the result of great magnification in the very shallow water around the edge. Accordingly the work done by the moon on this region may be neglected. The whole work done on the Yellow Sea by the moon is therefore small in comparison with the energy entering with the tide, and even its sign is uncertain. Thus the average dissipation in the whole of the sea is not very different from $1·1 \times 10^{18}$ ergs per second.

An alternative estimate may be deduced directly from the formula for the dissipation, with a suitable hypothesis on the distribution of velocity. At the entrance, when the current is flowing strongest, the dissipation in a strip a centimetre wide across it is $0{\cdot}002\rho\int u^3 dy$, and if the above distribution of velocity is correct this is $1{\cdot}2\times10^{11}$ ergs per second. Farther north the velocity is not so great; in fact, around the Shan-tung promontory it does not exceed one knot, and at the islands in Korea Bay opposite it is usually about two knots. Further, the width is here less than 300 km., so that on this account also the dissipation for a given velocity must be less; the amount in a strip a centimetre wide running east and west is therefore probably not more than one-sixteenth of that in a similar strip near the entrance. Even south of the narrow part off Shan-tung the currents appear to be slower than near the entrance; for about half the distance the velocity is about two knots, rising again to $3\frac{3}{4}$ knots at the Sir James Hall group, in the narrow region. If a proportional reduction takes place at all other distances from land, we must suppose that the above estimate of the dissipation per unit length is correct for the first 200 km., that the amount for the next 230 km. is an eighth of this, and that for the remaining 220 km., corresponding to Korea Bay, is a sixteenth. The Gulfs of Pe-Chili and Liau-tung may be ignored. The total dissipation, if the maximum velocity at every place occurred at the same time, would therefore be $(2\times10^7+\frac{1}{8}\times2{\cdot}3\times10^7+\frac{1}{16}\times2{\cdot}2\times10^7)\ 1{\cdot}2\times10^{11}$ ergs per second, or $2{\cdot}8\times10^{18}$ ergs per second. The average for each place is $4/3\pi$ of the maximum there, so that the average for the whole sea is $4/3\pi$ of this maximum, or $1{\cdot}2\times10^{18}$ ergs per second, which agrees with the previous estimate much more closely than the data would have led us to expect.

3. *The Sea of Japan.*

The Sea of Japan is an oval basin bounded on the eastern side by Japan and Sakhalin. It seems clear that the dissipation is small, for both the tide height and the current are small. Even in the comparatively narrow and shallow Korea Strait, through which the tide enters at the south end, the current only attains a speed of 1 or 2 knots; and when this opens into the sea the width suddenly increases to 900 km. and the depth to 400 fathoms. The currents in most of the sea must therefore be insignificant. They are appreciable at the gaps between the Japanese islands and in part of the narrow Gulf of Tartary, but the area affected is small and the currents only moderate (1 to 3 knots at most) so that the dissipation is small.

4. *The Sea of Okhotsk.*

In its essential features this resembles the Sea of Japan. The only shallow parts of it are narrow strips around the coast, while the tide enters through the shallow water of the straits between the Kurile Islands. As the tide in the sea depends on the supply of water to maintain it, the restriction on it imposed by the shallowness of the entrances causes the currents to be small. In the Gulfs of Ghijinsk and Penjinsk, in the north-east corner, the depth diminishes considerably, and the currents increase to 1½ to 2 knots. Data are very scanty, but the area affected appears to be about 70,000 sq. km., leading to a dissipation when the currents are strongest of from 6×10^{17} to $1 \cdot 4 \times 10^{18}$ ergs per second. If we adopt the mean of these as giving roughly the actual dissipation, and apply the factor $4/3\pi$, we find the average dissipation in these gulfs to be about 4×10^{17} ergs per second. There is probably no important dissipation elsewhere in the Sea of Okhotsk.

5. *The Bering Sea.*

In the extreme north of the Pacific, between Siberia and Alaska, a chain of small islands, the Aleutian Islands, extends all the way across. The region north of these has the shape of a quadrant and forms the Bering Sea. Between the islands the depth is great, and the tide of the Pacific seems to enter almost unhindered. Since the depth of more than half of the sea, mostly on the Alaskan side, is less than 40 fathoms, large currents are produced, especially in the three chief bays—the Gulf of Anadir, Norton Sound, and Bristol Bay. The dissipation must therefore be very

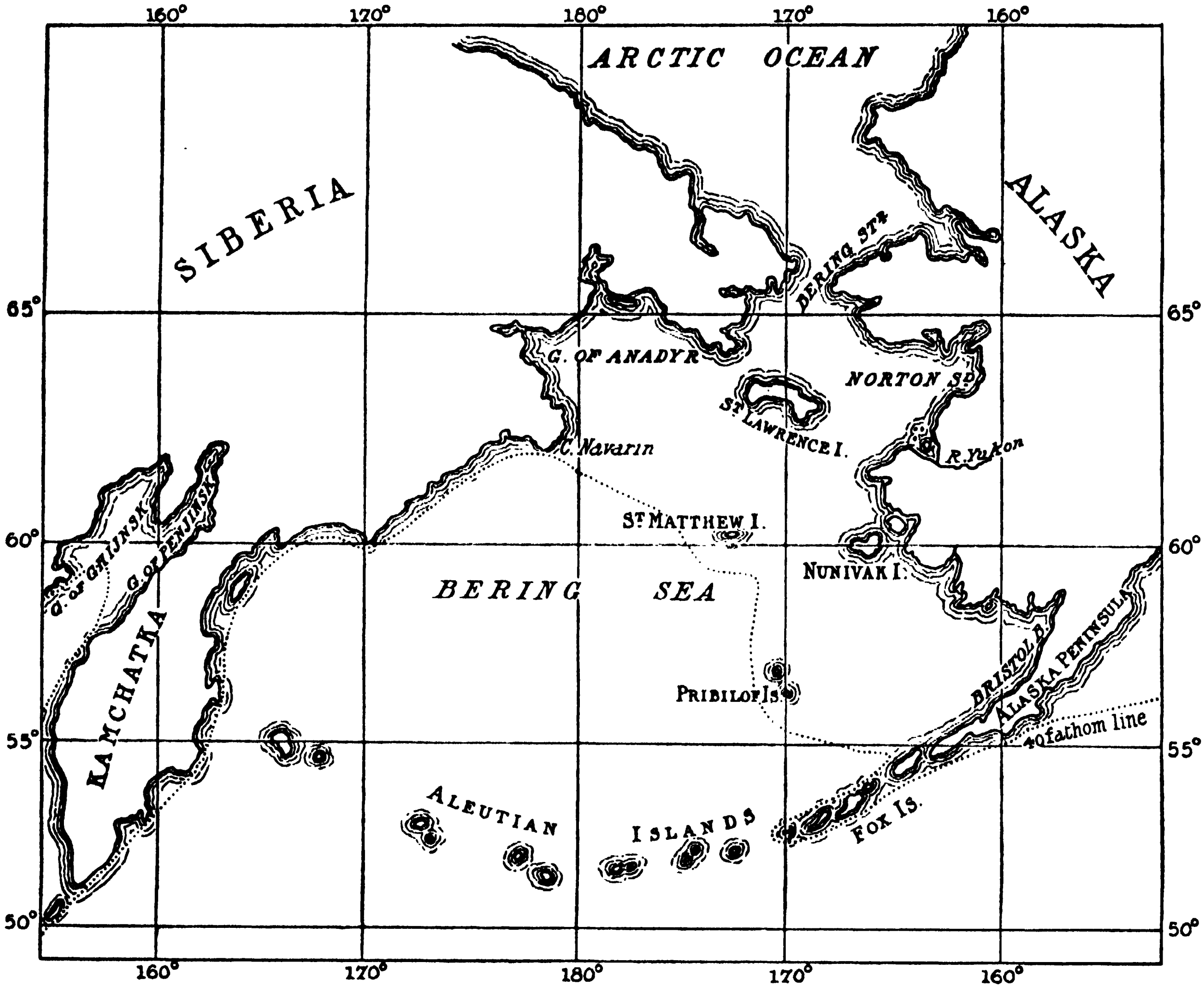


Fig. 3.

great, but a reasonably accurate estimate of it is difficult to make on account of the form of the shallow portion, which has no narrow place that can be called an entrance. It is best to treat the main part of the sea and the bays separately.

In the south of the sea it is stated that the maximum rate of the water, when clear of the passes between the Aleutian Islands, is usually about 2½ knots when the depth is less than 100 fathoms. In the region satisfying these conditions the depth is in most places about 80 fathoms, so that the current farther north, where the depth is often only 20 or 30 fathoms, may exceed this. On the other hand there seems to be little semi-diurnal tide in the extreme north. In Norton Sound the tide is diurnal, presumably because the waves from the south and from Bering Strait neutralize each other. At St. Lawrence Islands, near the entrance to the strait, the tide is only about a foot in height, confirming this suggestion. Farther south, however, the tidal wave from the Arctic must spread out and become inappreciable. At Pribilof Islands; 500 km. from the nearest land and surrounded by water 50 fathoms deep, the current reaches 2½ knots. At St. Matthew Island, about midway between these and St. Lawrence Islands, in water 30 to 40 fathoms deep, and nearly as far from land, the current still reaches 2½ knots. There are no other data given for islands far from shore, and it seems that we shall not be overestimating the dissipation if we take the maximum current to be 2½ knots all over the shallow region bounded on the south by the Fox Islands and extending north till half-way between St. Matthew and St. Lawrence Islands. The size of this is 1,000 by 700 km., or 7×10^5 sq. km. The maximum dissipation is therefore $2{\cdot}74 \times 10^{12} \times 7 \times 10^5 \times (2{\cdot}5)^3$, or 3×10^{19} ergs per second, and the mean dissipation $1{\cdot}2 \times 10^{19}$ ergs per second.

In Bristol Bay the average velocity seems to be about 3 knots, though the observations are few. The corresponding dissipation is about $1{\cdot}5 \times 10^{18}$ ergs per second. In Norton Sound the dissipation is probably small, for it is mostly north of St. Lawrence Island, and the tide is diurnal. The Gulf of Anadir probably contributes about as much as Bristol Bay; for though its area is twice as great, its more northerly situation must reduce the current somewhat. In all, then, the average rate of dissipation in Bering Sea is about $1{\cdot}5 \times 10^{19}$ ergs per second. This estimate is of course subject to considerable error, for it depends wholly on a few observations, which may not give quite a fair sample of the whole of the sea. The depths around the localities considered are fairly typical of the sea as a whole, so that great error on this ground is not to be anticipated; but errors in observing the velocities may be greater, and both kinds of error are magnified in importance by the fact that the velocity must be cubed when the dissipation is calculated, so that if the true mean velocity were only 2 knots instead of 2½ knots the dissipation would be almost halved. It does not appear that the velocity increases much towards the coast; in fact the velocities near the Alaskan coast seem to be rather smaller than those near the islands. Thus an underestimate on this ground is not probable.

6. *Malacca Strait.*

This is a narrow triangular area, about 800 km. in length, separating Sumatra from the Malay Peninsula. The tide of the Bay of Bengal enters at the north-west end, and gradually increases in height as it advances along the strait towards Singapore. At the south end, however, the part of the tide that has not been reflected or dissipated on the way through the strait is overwhelmed by the diurnal tide of the South China Sea. Ample observations of the tides and tidal currents on both sides are available. The currents as far south as Cape Medang (nearly due west of Malacca) seem to reach maxima of 1½ to 3 knots, the average amplitude at springs being practically 2 knots. The area of this region is 100,000 sq. km., and the dissipation is accordingly found to be about 9×10^{17} ergs per second on an average.

An alternative determination can be made by finding the rate of inflow of energy. At Kumpei, on the Sumatran side and near the north end of the strait, it is high water, full and change, at noon, and the amplitude at springs is 120 cm. The flood tide outside the bar, in water about 20 fathoms deep, sets south-east from three hours before high water till three hours after it, so that it reaches its maximum speed of 1½ knots at high water. At Penang, near the opposite shore, it is high water at 0h. 21m. and the current reaches its maximum velocity of 2½ knots an hour before high water. As the strait is everywhere narrow in comparison to its length, and as these observations do not seem to have been taken on shoals, they are probably representative of that part of the strait. The amplitude of the tide at Penang is 100 cm. We can therefore take the average height of the tide along the section from Kumpei to Penang to be 110 cm., and the average current when flowing strongest to be 2 knots, reaching its maximum half an hour before high water. The average depth is about 30 fathoms. A modification must be made in the previous procedure to allow for the fact that the current flows along the strait, which is not quite at right angles to the line of the section; therefore, in finding the energy crossing the section, we must take for the length of the section, not the distance from Kumpei to Penang, but the projection of this on a line perpendicular to the strait, which is 230 km. The flux of energy is hence found to be, on an average, 7×10^{17} ergs per second. We also require the amount of this energy that emerges through the narrow part of the strait. Off Cape Medang, which marks the narrowest point of the strait away from the immediate neighbourhood of Singapore, the amplitude of the tide is 120 cm., high water occurring at 6h. 30m. The tidal current flows at an average speed of about 2½ knots. There is no record of the tidal phenomena just opposite, but in Malacca Road it is high water at 7h. 30m., with an amplitude of 165 cm.; the current there reaches its maximum of 2 knots an hour before high water. At the eastern end of South Sands, which lies near the Malay side, north-west of Cape Medang, it is high water about 6h. 0m., and the tidal stream has a maximum speed of 1½ knots an hour before high water. The width of the channel at Cape Medang is

36 km., and the depth 20 fathoms. The average flux of energy eastward past it is found to be $1{\cdot}0 \times 10^{17}$ ergs per second. Thus the average excess of the inflowing energy over the issuing energy is 6×10^{17} ergs per second. The work done by the moon is insignificant, for it crosses the meridian at full and change at 2h. 30m., which is nearly the average time of high water. Thus all the excess of energy just found is dissipated in the strait.

The area of the strait between the Kumpei-Penang section and the Medang section is 56,000 sq. km. If the average current in this had an amplitude of 2 knots the dissipation would be 5×10^{17} ergs per second on an average, in striking agreement with the estimate from the flux of energy, though the latter is more reliable. If in the final estimate the region north of the Kumpei-Penang section is to be included, we must add a fraction to the total to allow for it, making probably between 8×10^{17} and 12×10^{17} ergs per second in all.

In the part of the strait east of Medang there are few records of the currents, but the dissipation is probably small. In any case it could not exceed the 10^{17} ergs per second that pass Medang, and is probably less than this. The total dissipation in the Strait of Malacca is therefore $1{\cdot}1 \times 10^{18}$ ergs per second, subject to an uncertainty of a fifth of its amount.

AUSTRALIAN WATERS.

Australia is surrounded by a belt of water less than 100 fathoms in depth ; the width of this ranges from 10 to 200 miles, except at the Gulf of Carpentaria, where it extends right across to New Guinea. The tide in this neighbourhood is diurnal, like that in the South China Sea to the north of it. The contribution to the secular acceleration of the moon is accordingly very small. The tidal streams do not exceed 1 knot, and as the area is much less than that of the South China Sea the dissipation in the diurnal tide cannot be comparable with that already found for the larger sea.

AFRICAN WATERS.

The Mozambique Channel.

The channel between Madagascar and the mainland is mostly about 500 fathoms deep or more. There are few records of tidal currents in it; in fact the only record given in the 'African Pilot,' part 3, appears to be based on the statement of a single observer, that the tidal streams in the channel are comparable with the permanent current driven by the trade winds, which flows at about 2 knots. This cannot however be uniform all over the channel, for the following reason. The height of the tide along the African coast is about 12 feet, which is as usual measured relative to low water at ordinary springs, so that the vertical amplitude of the tide is 180 cm. Now the ordinary theory of tides in channels shows that the maximum velocity is of order $h\,(g/\mathrm{D})^{\frac{1}{2}}$; in a simple wave in a uniform channel it is exactly this. Taking the depth to be 90,000 cm., this makes the maximum velocity 19 cm./sec., or rather more than a third of a knot. Accordingly the currents with velocities of a knot or more must be confined to narrow coastal strips, and the dissipation is therefore small.

The only other partially enclosed regions around Africa are the Gulf of Aden and the Red Sea. The former is deep in the middle with narrow strips of shallow water on the margins, like the Mozambique Channel, and therefore the dissipation is small. The Red Sea is shallower, but can have no important currents, since the inlet at Aden is so narrow. The Mediterranean has already been dealt with. Thus the dissipation around the coasts of Africa is negligible.

North American Waters.

There are many partially enclosed bodies of water around North America, the chief of which are the Gulfs of Mexico and California, the Bay of Fundy, the Gulf of St. Lawrence, and the numerous straits and bays of the North-west Passage. Of these the Gulf of Mexico may be ruled out at once, for it is very deep and a large fraction of its entrance is blocked by Cuba. The Gulf of California is still deeper; and therefore the currents in these cannot be notable except in restricted localities.

1. *The Bay of Fundy.*

This bay requires to be considered separately in spite of its small size, for it is famous for possessing the largest tides in the world. It is fairly shallow, and the tides are much magnified in height by the diminution in both depth and width towards the head of the Bay. The currents are apparently not so great as would be expected from the height of the tides. The entrance is through the Grand Manan Channel, named after an island in it. The average current in the channel reaches about 1·8 knots, and that near St. John, half-way up the bay, reaches 1·7 knots. The area of the bay is 12,000 sq. km., so that the average dissipation for a maximum velocity of 1·8 knots all over would be $7{\cdot}7 \times 10^{16}$ ergs per second. It is likely that the currents farther up the bay are stronger, so that this must be regarded as a lower limit.

An alternative estimate may be obtained from the inflow of energy. The rise of the tide in Grand Manan Channel is at most places about 20 or 22 feet at springs, above low water ordinary springs. The amplitude is therefore about 320 cm. The current has an amplitude of 1·8 knots, and the depth of the channel is about 9000 cm. The average time of the turn of the current at three places near the south side of the channel (those numbered 14, 16 and 17, in the 'Nova Scotia and Bay of Fundy Pilot,' page 22) is 35 minutes after high water at St. John. This high water at full and change occurs at 11h. 21m., while at l'Etang, on the north side of the channel, it is at 11h. 18m., and the mean of the times at Westport, Petit Passage, and Digby Gut, which are nearly opposite, is 10h. 48m. Thus the mean time of high water in the channel must be about 11h. 3m., so that the current turns 53 minutes after the tide. The phase difference is therefore 25 degrees. The width of the channel is 83 km. Applying equation 15 of TAYLOR's paper, we find that the average rate at which energy enters is $4{\cdot}7 \times 10^{17}$ ergs per second.

The average rate at which the moon does work on the bay is $-\frac{1}{2}g \int hh'\gamma \sin \beta dS$, as was found for the Yellow Sea. In this case the latitude is 45° north, so that h' is 18 cm. In the lower half of the bay the amplitude of the tide is not greater than 12 feet, but in the upper half it rapidly increases, till in Minas Basin it reaches 25 feet and in Chignecto Bay 23 feet. The time of high water in the bay ranges from 11h. 3m. to 11h. 50m. The later value corresponds to the upper part, where the amplitude is greatest; but as this part is also the narrowest, the two times must receive about equal weights in finding the average. We therefore take the average time of high water to be 11h. 30m. The average amplitude of the tide is about 18 feet, or 540 cm. The longitude of the bay is 66° west, so that the moon crosses the meridian at full and change at 4h. 33m. The time of high water is more than 6h. 12m. later than this, so that the tide is falling when the moon is exerting its greatest upward pull, and the work done by the moon is therefore negative. The interval between transit and low water is 45 minutes, so that $\beta = 22°$. The area of the bay is $1{\cdot}2 \times 10^{14}$ sq. cm. The work done by the moon is therefore -3×10^{16} ergs per second. The total dissipation in the bay is $4{\cdot}4 \times 40^{17}$ ergs per second.

This estimate is six times as great as the earlier one based on the currents alone. It is much the more reliable, for the first depended on the assumption that the currents were equally great all the way up the bay, whereas actually they increase very much towards the head. Velocities up to 9 knots are recorded in Minas Basin, though the area in which these occur must be very restricted. The most serious source of error in the second estimate is the phase difference, for this is only an hour and would be affected to a considerable extent by an error in the determination of the time when the current turns. The observed time of turn does not vary much from place to place, however, and it does not seem likely that the estimate is wrong by more than a quarter of its amount. The second estimate will therefore be adopted. It will be noticed that it is rather less than the dissipation in the Irish Sea.

The Gulf of St. Lawrence gives very little dissipation. The narrow entrance through Cabot Strait prevents the tides from being considerable except in the estuary of the river itself and in Belle Island Strait, which separates Newfoundland from Labrador.

2. *The North-west Passage.*

The channel from the Atlantic to the Arctic between Canada and Greenland is blocked by a large number of islands of varying sizes, between which are narrow and shallow straits. The dissipation in several of these can be estimated from data in the 'Arctic Pilot,' vol. 3.

The chief of these channels is Davis Strait, with Baffin Bay to the north of it, which lies between Baffin Land and Greenland. It is about 1600 km. in length. The only tidal velocities recorded in it, except in fjords, are near Holstenborg, in Greenland, where the currents in the offing are said to reach a speed of two knots. This cannot, however, be general, for there is a shallow region off Holstenborg, some 150 km. long and 60 km. wide, with a depth of about 23 fathoms. Most of the strait is about 100 fathoms deep. This region is therefore a place where the main current of the strait is magnified by the form of the bottom, and there is no reason to believe that the current in the deep water is greater than half a knot, which is the observed velocity off the coast of Labrador. The dissipation in Davis Strait is therefore not great.

At the northern extremity of the strait there are several narrow passages into the Arctic. The dissipation in these must be important. In Smith Sound and Kennedy Channel, for instance, which separate the north-west coast of Greenland from Ellesmere Land, the current is said to be "nearer two figures than one." These straits are small in area, but if such currents exist over much of their extent we must take them into account. The data available at present are unfortunately too meagre.

The south end of Davis Strait is connected to Hudson Bay by Hudson Strait. The currents in this are described as "great enough to be dangerous," especially at the east end; but the recorded currents, even in the middle of the strait, are only about three-quarters of a knot. This makes the average dissipation about 5×10^{16} ergs per second. The danger arises mostly from drifting ice.

In the entrance to Hudson Bay the velocity increases to one and a half knots. The area over which this is true is about $3{\cdot}8 \times 10^{14}$ sq. cm., making the average dissipation $1{\cdot}5 \times 10^{17}$ ergs per second.

In Hudson Bay itself the currents are probably very small. Considerable velocities are recorded at Port Churchill, but there are no records in the middle of the bay. The depth in the middle is about 50 fathoms, which is about the same as at the entrance. The entering current must therefore spread out in the bay and undergo great diminution in strength. Near Port Churchill the depth is only 19 fathoms or less, so that the current there must be a local current magnified. The dissipation in Hudson Bay must therefore be small.

In Fox Strait, which runs northwards from the entrance to Hudson Bay, the depth is less, about 20 fathoms, and the current reaches one and a half knots. The area of this channel is 2×10^{15} sq. cm., and the appropriate average dissipation is about $1{\cdot}4 \times 10^{18}$ ergs per second. The remaining straits of the North-west Passage probably do not contribute nearly so much to the dissipation, for the energy of the entering wave must be mostly dissipated in the channels already dealt with, and partly through this and partly on account of the obstructive effect of the islands it is not likely that the straits farther north-west are very important, though this cannot be regarded as certain. The dissipation in the whole of the North-west Passage is thus about $1{\cdot}6 \times 10^{18}$ ergs per second on an average. Adding this to the amount found for the Bay of Fundy, we have for the whole of North America a total of 2×10^{18} ergs per second on an average.

SUMMARY.

The mean rates of dissipation in the lunar semi-diurnal tide found in the foregoing investigation are as follows:—

	Ergs per Second.
European waters	$2{\cdot}4 \times 10^{18}$
Asiatic waters:	
South China Sea	Small.
Yellow Sea	$1{\cdot}1 \times 10^{18}$
Sea of Okhotsk	$0{\cdot}4 \times 10^{18}$
Bering Sea	$15{\cdot}0 \times 10^{18}$
Malacca Strait	$1{\cdot}1 \times 10^{18}$
North American Waters:	
Bay of Fundy	$0{\cdot}4 \times 10^{18}$
North-west Passage	$1{\cdot}6 \times 10^{18}$

The total thus accounted for is $2{\cdot}2 \times 10^{19}$ ergs per second. I have shown in a previous paper that the dissipation required to account for the secular acceleration of the moon (which amounts to 9″ per century per century) is about $1{\cdot}4 \times 10^{19}$ ergs per second, so that it seems as if there is more dissipation than is required. If this was so it would be necessary to seek for a cause that could produce an appreciable secular retardation of the moon, and none such is known. A scrutiny of the results so far obtained is therefore desirable, with a view to finding out whether any of them have been over-estimated. One cause of such an over-estimate is easily seen. The data used for the Irish Sea, the English Channel, Malacca Strait and the Bay of Fundy refer definitely to spring tides alone when the currents are at a maximum. The height of the tide adopted in the calculation for the Yellow Sea was also that of the spring tide. In the other cases it is not stated whether the currents have average or spring values, but if they were determined at springs the requisite reduction is at once obtained. The theoretical ratio of the heights of the lunar and solar tides is 2·3 when inertial and frictional effects are neglected. This ratio is probably nearly correct in mid-ocean, for the periods of the two tides are not very different, so that inertia will affect their amplitudes in the same ratio. In shallow areas, however, the frictional force is not proportional to the velocity but to its square, and accordingly friction has more relative effect in reducing the tides when they are great than when they are smaller. The ratio of the solar to the lunar tide, as found from observations, is accordingly less than the theoretical value, since the ratio of the ranges at springs and neaps is reduced. This ratio is stated in the 'Admiralty Tide Tables for 1920' to be 1 : 2·73 on an average. If now θ be the phase of the lunar tide, let $(1-r)\theta$ be the phase of the solar tide, so that r is 1/29. Let A be the amplitude of the lunar tidal current and $A\nu$ that of the solar tidal current. The total current is

$$A\{\cos\theta+\nu\cos(1-r)\theta\} = A\,(1+2\nu\cos r\,\theta+\nu^2)^{\frac{1}{2}}\cos\left(\theta-\tan^{-1}\frac{\sin r\,\theta}{1+\nu\cos r\,\theta}\right)$$

which is now expressed as a simple harmonic motion with a slowly varying amplitude and period. The amplitude at springs is $A(1+\nu)$. The dissipation is proportional to the cube of the current, and therefore to the cube of the amplitude. The ratio of the mean dissipation to the dissipation at springs is therefore the average of $(1+2\nu\cos r\theta+\nu^2)^{\frac{3}{2}}/(1+\nu)^3$. If ν^6 be neglected, the numerator of this is $1+\frac{9}{4}\nu^2+\frac{9}{64}\nu^4$. Assuming that the ratio of the velocities is the same as that of the vertical ranges, we find that this fraction is equal to 0·51. Applying this correction to the spring tide dissipation, we find that the average dissipation is $1{\cdot}1\times10^{19}$ ergs per second, 80 per cent. of what is required. It would give a secular acceleration of the moon of 7″ per century per century.

The agreement between the dissipation in shallow seas and that necessary to account for the lunar secular acceleration is much closer than the data would entitle us to expect. Two-thirds of that found takes place in the Bering Sea, the estimate for which may be incorrect by half its amount. What we are entitled to assert, however, is that this dissipation is certainly enough to account for a large fraction of the secular acceleration, and that there is nothing to prove that it is incapable of accounting for the whole of it.

It is uncertain whether the dissipation in any other coastal regions is notable in comparison with those already considered. The only partly enclosed areas not treated here that are of considerable size are some of those in the North-west Passage. There is an extensive shallow region off the coast of Patagonia, but it is in no way enclosed, being perfectly open to the Atlantic. Thus it is difficult to make any reliable inference about the currents in it. Many records of tidal currents along the coast are given, some reaching several knots, but all of them seem to refer to currents up rivers or near their mouths, where the general currents must be magnified, or to currents over bars and shoals; there seem to be no data about the currents more than a few miles out to sea.

The dissipation over local shallows like shoals and bars and in narrow bays and straits has been systematically ignored in this paper, except where it has been automatically taken into account in the determination of the excess of the entering over the issuing energy. The chief reason for this is the utter impossibility of finding it. The fjords on the west coasts of Norway, Greenland, and North and South America are innumerable, and in many of them, perhaps in all, there is a strong tidal current, so that the dissipation per unit area in these places must very much exceed that in any of the areas here treated. On the other hand, the total area must be less, and it is uncertain whether the increase in velocity is enough to counterbalance the decrease in area and make the total dissipation in these places comparable with that here found for the larger shallow seas. The same is true of shoals ; though the agreement between the results given by the two methods of finding the dissipation in shoaly waters, as in the Yellow Sea and the Strait of Malacca, indicates that the shoals at any rate do not contribute to the dissipation an amount overwhelmingly greater than the normal places, for one method necessarily includes the effect of the shoals and the other systematically omits it. Along the open shore again there must be some dissipation ; the currents there do not usually extend many miles out to sea, but they exist along a very long stretch of coast, and the aggregate dissipation in them may be appreciable.

The hypothesis that the secular acceleration of the moon is due to dissipation of energy in the tides in shallow coastal regions therefore seems capable of satisfying all the quantitative demands on it, and it is also free from objections that have been urged against other attempted explanations.* It therefore occupies a strong position.

APPENDIX.

The Secular Change in the Obliquity of the Ecliptic.

In consequence of the dissipation of energy in the diurnal tides there must be a couple always acting on the earth so as to tend to resist its angular motion about an axis in the plane of the orbit of the moon or the sun, as the case may be. If δ be the declination of the moon, the angular velocity of the earth about the diameter that points to the moon is $\Omega \sin \delta$, and the angular momentum about it is $C\Omega \sin \delta$, where C is the earth's moment of inertia. Let L be the couple about this diameter. Then the rate of dissipation of energy in the diurnal tide is $L\Omega \sin \delta$. Also the rate of change in the inclination is given by

$$\frac{d}{dt}(C\Omega \sin \delta) = L.$$

Now L must contain $\Omega \sin \delta$ as a factor, since it depends for its existence on the existence of the diurnal tide, whose coefficient is proportional to $\sin \delta$, and whose speed is equal to $\Omega - n$, which is sufficiently near to Ω for our present purpose. Hence the friction of the tide will have a damping effect on this component of the earth's rotation, which is not altered by the other couples acting, since none of these have a notable effect in diminishing the amplitude of the motion. If the average value of the dissipation is taken to be 5×10^{18} ergs per second, this being rather more than was found in the South China Sea, and we remember that the average of $\sin^2 \delta$ is $\frac{1}{2} \sin^2 i$, where i is the obliquity of the ecliptic, we find that the amplitude of $\Omega \sin \delta$ would be reduced to $1/e$ of its value in 2×10^9 years. This is of the same order as the probable age of the earth. If Ω remained constant this would show that the inclination of the Equator to the ecliptic would be reduced by $1''$ in about 2×10^4 years. Actually Ω is decreasing, so that if the rate were maintained it would be reduced to $1/e$ of its value in about 10^{10} years, a longer time than was found, on the assumption stated, to be enough for a similar reduction in the obliquity. Thus we can infer that the obliquity is at present diminishing, though there is no reason to believe that there has been any observable change in it in historic times. Even if there were as much dissipation in the diurnal tides as in the semi-diurnal ones this would hardly be possible.

---

* *Cf.* 'Monthly Notices of R.A.S.,' vol. lxxx., 1920, pp. 309–317.

[*Note added September 16.*—Mr. TAYLOR asks me to point out certain errata in his paper "Tidal Friction in the Irish Sea." On p. 2, $1 - \frac{\gamma}{\sqrt{r}}$ is twice written for $1 + \frac{\gamma}{\sqrt{r}}$; the correct form is used in equations (4) and (5). On p. 9, line 9, T_1 is Greenwich mean time of high water at full and change of the moon at the place considered, whereas the "establishment" is the *local* mean time of this event. In equation (16), $t + T_1$ should be $t - T_1$, and in equation (18), $t + T_0$ should be $t - T_0$. On pp. 19 and 20, $\sin^2 \phi_0$ is consistently written for $\sin 2\phi_0$; equation (33) is correct.

In this paper, as in TAYLOR'S, integrals over a period are always determined as if the current velocity and the tide height varied harmonically. This could be strictly correct only if the frictional force was proportional to the velocity, which is not the case. It appears, however, that the departure from the harmonic variation is not enough to produce any great alteration in these integrals.

I wish to express my thanks to Mr. H. W. BRABY for drawing the maps in this paper.]

IX. *Plane Stress and Plane Strain in Bipolar Co-ordinates.*

By G. B. JEFFERY, *M.A.*, *D.Sc.*, *Fellow of University College, London.*

Communicated by Prof. L. N. G. FILON, *F.R.S.*

Received May 15,—Read June 24, 1920.

§ 1. INTRODUCTION.

THE problem of the equilibrium of an elastic solid under given applied forces is one of great difficulty and one which has attracted the attention of most of the great applied Mathematicians since the time of EULER. Unlike the kindred problems of hydrodynamics and electrostatics, it seems to be a branch of mathematical physics in which knowledge comes by the patient accumulation of special solutions rather than by the establishment of great general propositions. Nevertheless, the many and varied applications of this subject to practical affairs make it very desirable that these special solutions should be investigated, not only because of their intrinsic importance but also for the light which they often throw on the general problem. One of the most powerful methods of the mathematical physicist in the face of recalcitrant differential equations is to simplify his problem by reducing it to two dimensions. This simplification can only imperfectly be reproduced in the Nature of our three-dimensional world, but, in default of more general methods, it provides an invaluable weapon.

It was shown by AIRY* that in the two-dimensional case the stresses may be derived by partial differentiations from a single stress function, and it was shown later† that, in the absence of body forces, this stress function satisfies the linear partial differential equation of the fourth order $\nabla^4\chi = 0$, where $\nabla^4 = \nabla^2 . \nabla^2$, and ∇^2 is the two-dimensional Laplacian $\partial^2/\partial x^2 + \partial^2/\partial y^2$.

It might have been expected that these results would have opened the way for a theory of two-dimensional elasticity of the same generality as the two-dimensional potential theory. This has not, however, been the case. This is due in part to the greater analytical difficulties which attend the discussion of the two-dimensional solutions of $\nabla^4\chi = 0$ as compared with $\nabla^2\chi = 0$. The analogues of many of the important properties of the simpler equation have yet to be discovered if they exist at all. Some progress has been made, and in this connection we may mention the work of J. H. MICHELL[*] who established a general theory of inversion which, with some important differences, follows the potential theory fairly closely.

---

* 'Brit. Assoc. Rep.,' 1862, p. 82.

† W. J. IBBETSON, 'Proc. Lond. Math. Soc.,' vol. xvii., 1886, p. 296. For a history of this part of the subject see LOVE'S 'Elasticity,' 2nd edition, p. 17.

No doubt the analytical difficulties have been the chief obstacle to progress, but perhaps the theory has not in recent years received attention which it would have received but for a certain physical difficulty. A truly two-dimensional elastic system is not so easy of realisation as might seem to be the case at first sight. If the stresses are everywhere parallel to the xy plane and independent of z there will in general be a varying displacement parallel to z. If the displacements are everywhere parallel to the xy plane and independent of z this can only be secured by the application of a stress $\widehat{zz}$ which varies from point to point and is perpendicular to the xy plane. This difficulty was in a large measure removed by a theorem established by FILON, which has been called the theorem of generalised plane stress.[†] It states that if the average value of the stress $\widehat{zz}$ be taken throughout the thickness of a plate parallel to the xy plane, then the ordinary two-dimensional theory will give accurately the *average* stresses through the thickness of the plate if the elastic constants of the material are modified. If λ, μ denote the true elastic constants, λ must be replaced by $\lambda' = 2\lambda\mu/(\lambda+2\mu)$ while μ remains the same as before. This theorem attains an even greater importance when considered in the light of MICHELL'S theorem,[‡] that if a plate bounded by any number of bounding curves is in equilibrium under forces in its plane applied over the boundaries, then, provided the forces applied over each boundary taken separately are in equilibrium, the stresses are everywhere independent of the elastic constants. The hypothesis that the average value of $\widehat{zz}$ vanishes throughout the plate, while certainly not accurately true in the majority of cases, will probably give a very close approximation in the case of a *thin* plate where parallel faces are unstressed.

In the light of this generalisation it is of considerable importance that the two-dimensional problem should be worked out more thoroughly. The two-dimensional solutions of $\nabla^4\chi = 0$ have been investigated in several systems of curvilinear co-ordinates. Owing to the special importance of the problem of the rectangular beam the solutions in Cartesian co-ordinates have naturally received a considerable amount of attention. MICHELL gave the general form of the stress-function in polar co-ordinates, thus opening the way for the solution of the problem of a plate bounded by two concentric circles, or an infinite plate containing a circular hole under any given tractions applied over its boundaries. In his lectures at University College, London, in 1912, Prof. FILON gave the complete solution of this problem determining the stresses and displacements when the stresses on the boundaries are expanded in Fourier series, and I am not aware that this solution has ever been published. An outline of the solution in elliptic co-ordinates is given in LOVE's 'Elasticity.'*

---

[*] "The Inversion of Plane Stress," 'Proc. Lond. Math. Soc.,' 1901, vol. xxxiv., p. 134. Many of the results of the present paper can be obtained by an application of MICHELL'S methods, but it has proved more convenient to proceed on different lines.

[†] 'Roy. Soc. Phil. Trans.,' A, 1903, vol. 201, pp. 63–155.

[‡] 'Proc. Lond. Math. Soc.,' vol. xxi., 1900, p. 100.

In this paper the complete solution is given for bipolar co-ordinates, for which the co-ordinate curves are co-axial circles. This solution enables us to treat the problems of an infinite plate containing two circular holes, a semi-infinite plate bounded by a straight edge and containing one circular hole, and a circular disc with an eccentric circular hole.

In the second Section the equations are expressed in bipolar co-ordinates and formulæ are established for the displacements in terms of the stress-function.

In the third Section the stress-function is obtained in a convenient form and the terms giving rise to many valued displacements are separated out.

The fourth Section is devoted to the determination of the coefficients in the stress-function when the tractions over the boundaries are given in Fourier series, and to an examination of the convergence of the resulting series. From the results established in this section it appears that the solution is complete, for the stress-function can always be uniquely determined when the tractions are given, provided that the applied forces taken as a whole are in equilibrium.

The remaining sections are occupied with the examination of some of the simpler applications of the theory. Section 5 gives the solution for a circular disc with an eccentric hole (or a cylinder with eccentric bore) when the two boundaries are under different hydrostatic pressures. It is found that the solution of this problem can be expressed in finite terms. An important particular case of this problem is discussed in Section 6, namely, a semi-infinite plate with a straight, unstressed boundary and a circular hole under a uniform normal pressure. This will give the stresses near a rivet hole while the hot plastic rivet is being forced home under pressure. This solution is interesting from another point of view, for if the ratio of the radius of the hole to its distance from the edge is suitably adjusted, the point of greatest tension will be on the straight edge while the point of greatest stress difference is on the circular boundary. It thus suggests a crucial test for the rival theories of rupture,—the greatest tension theory and the greatest stress-difference theory.

Section 7 deals with a semi-infinite plate with an unstressed circular hole under tension parallel to its straight edge. The solutions are in the form of infinite series, but the more important aspects of the problem are illustrated by numerical tables.

---

* 2nd edition, p. 259.

§ 2. THE CO-ORDINATES.

Let us take curvilinear co-ordinates defined by the conjugate functions

$$\alpha+i\beta=\log\frac{x+i(y+a)}{x+i(y-a)}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (1)$$

where x, y are Cartesian co-ordinates and a is a positive real length. Solving for x, y, we have

$$x=\frac{a\sin\beta}{\cosh\alpha-\cos\beta}, \qquad y=\frac{a\sinh\alpha}{\cosh\alpha-\cos\beta}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

Elements of arc measured along the normals to the curves α, β = *constant* are respectively $\delta\alpha/h$, $\delta\beta/h$, where

$$\frac{1}{h^2}=\left(\frac{\partial x}{\partial\alpha}\right)^2+\left(\frac{\partial y}{\partial\alpha}\right)^2,$$

from which we have

$$h=(\cosh\alpha-\cos\beta)/a. \quad . \quad . \quad . \quad (3)$$

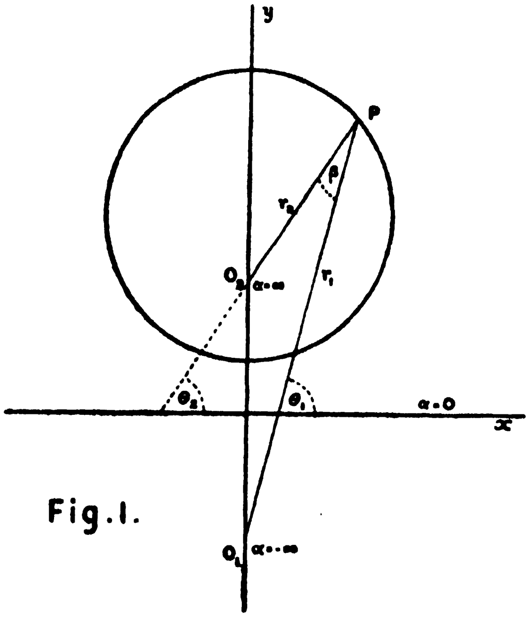


Fig. 1.

The general scheme of co-ordinates is shown in fig. 1. If O_1, O_2 are the points $0, -a$ and $0, a$ respectively and P any point in the plane, and if the radii from O_1, O_2 to P are of lengths r_1, r_2 and are inclined at angles θ_1, θ_2 to the axis of x, then $\alpha=\log r_1/r_2$ and $\beta=\theta_1-\theta_2$. The curves α = *constant* are a set of co-axial circles having O_1, O_2 for limiting points. The circles corresponding to positive values of α lie above the x-axis and those corresponding to negative values below, while the x-axis itself, which is the common radical axis, is given by $\alpha=0$. The curves β = *constant* are circles, or rather arcs of circles passing through O_1, O_2 and cutting the first set of circles orthogonally. On the right-hand side of the y-axis β is positive and on the left-hand side negative, while on the y-axis $\beta=0$, except on the segment O_1O_2, where $\beta=\pm\pi$. At infinity $\alpha=0$, $\beta=0$, and at O_1, O_2 we have $\alpha=-\infty$ and $+\infty$ respectively.

We have thus a set of co-ordinates adapted for the consideration of two-dimensional problems in which the region considered is—

(1) A finite region bounded internally by a circle and externally by a larger and non-concentric circle.

(2) A semi-infinite region bounded externally by a straight line and containing a circular hole.

(3) An infinite region containing two circular holes of any radii and centre distance.

If the displacements in the directions normal to the curves α and β constant are u, v respectively, the strains are given by*

$$e_{\alpha\alpha} = h\frac{\partial u}{\partial \alpha} - v\frac{\partial h}{\partial \beta}, \qquad e_{\beta\beta} = h\frac{\partial v}{\partial \beta} - u\frac{\partial h}{\partial \alpha},$$

$$e_{\alpha\beta} = \frac{\partial}{\partial \alpha}(hv) + \frac{\partial}{\partial \beta}(hu),$$

and the corresponding components of stress by

$$\left.\begin{aligned} \widehat{\alpha\alpha} &= \lambda\,(e_{\alpha\alpha}+e_{\beta\beta}) + 2\mu e_{\alpha\alpha}, \\ \widehat{\beta\beta} &= \lambda\,(e_{\alpha\alpha}+e_{\beta\beta}) + 2\mu e_{\beta\beta}, \\ \widehat{\alpha\beta} &= \mu e_{\alpha\beta}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (4)$$

These stresses may be derived from a stress-function, so that in rectangular co-ordinates

$$\widehat{xx} = \frac{\partial^2\chi}{\partial y^2}, \qquad \widehat{xy} = -\frac{\partial^2\chi}{\partial x\,\partial y}, \qquad \widehat{yy} = \frac{\partial^2\chi}{\partial x^2}.$$

Transforming these equations to curvilinear co-ordinates we obtain

$$\left.\begin{aligned} \widehat{\alpha\alpha} &= h\frac{\partial}{\partial \beta}\left(h\frac{\partial\chi}{\partial \beta}\right) - h\frac{\partial h}{\partial \alpha}\frac{\partial \chi}{\partial \alpha}, \\ \widehat{\beta\beta} &= h\frac{\partial}{\partial \alpha}\left(h\frac{\partial\chi}{\partial \alpha}\right) - h\frac{\partial h}{\partial \beta}\frac{\partial \chi}{\partial \beta}, \\ \widehat{\alpha\beta} &= -h\frac{\partial^2(h\chi)}{\partial \alpha\,\partial \beta} + h\chi\frac{\partial^2 h}{\partial \alpha\,\partial \beta}. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (5)$$

We will usually find it convenient to deal with $h\chi$ instead of χ itself, and in our particular co-ordinates these equations become

$$\left.\begin{aligned} a\widehat{\alpha\alpha} &= \left\{(\cosh\alpha - \cos\beta)\frac{\partial^2}{\partial \beta^2} - \sinh\alpha\frac{\partial}{\partial \alpha} - \sin\beta\frac{\partial}{\partial \beta} + \cosh\alpha\right\}(h\chi), \\ a\widehat{\beta\beta} &= \left\{(\cosh\alpha - \cos\beta)\frac{\partial^2}{\partial \alpha^2} - \sinh\alpha\frac{\partial}{\partial \alpha} - \sin\beta\frac{\partial}{\partial \beta} + \cosh\beta\right\}(h\chi), \\ a\widehat{\alpha\beta} &= -(\cosh\alpha - \cos\beta)\frac{\partial^2(h\chi)}{\partial \alpha\,\partial \beta}. \end{aligned}\right\} \quad . \quad . \quad . \quad (6)$$

We may note that

$$a\,(\widehat{\alpha\alpha} - \widehat{\beta\beta}) = (\cosh\alpha - \cos\beta)\left(\frac{\partial^2}{\partial \beta^2} - \frac{\partial^2}{\partial \alpha^2} + 1\right)(h\chi), \quad . \quad . \quad . \quad . \quad . \quad (7)$$

---

* LOVE, 'Theory of Elasticity,' p. 54.

so that if $h\chi$ and its second differential coefficients are finite at infinity ($\alpha = 0$, $\beta = 0$) we have there $\widehat{\alpha\alpha} = \widehat{\beta\beta} = h\chi$ and $\widehat{\alpha\beta} = 0$.

In the absence of body forces the stress-function satisfies $\nabla^4\chi = 0$. In curvilinear co-ordinates we have

$$\nabla^2 \equiv h^2\left(\frac{\partial^2}{\partial\alpha^2} + \frac{\partial^2}{\partial\beta^2}\right),$$

and, taking $h\chi$ as the dependent variable, we have in our co-ordinates

$$a\nabla^2\chi = \left\{(\cosh\alpha - \cos\beta)\left(\frac{\partial^2}{\partial\alpha^2} + \frac{\partial^2}{\partial\beta^2}\right) - 2\sinh\alpha\frac{\partial}{\partial\alpha} - 2\sin\beta\frac{\partial}{\partial\beta} + \cosh\alpha + \cos\beta\right\}(h\chi).$$

Repeating the operator, a little reduction leads us to the following transformation for $\nabla^4\chi = 0$:

$$\left(\frac{\partial^4}{\partial\alpha^4} + 2\frac{\partial^4}{\partial\alpha^2\partial\beta^2} + \frac{\partial^4}{\partial\beta^4} - 2\frac{\partial^2}{\partial\alpha^2} + 2\frac{\partial^2}{\partial\beta^2} + 1\right)(h\chi) = 0. \quad . \quad . \quad . \quad . \quad . \quad (8)$$

Thus by considering $h\chi$ instead of χ we have a linear equation with constant coefficients.

Before proceeding to the discussion of its solutions, we must investigate the method of determining the displacements corresponding to a given stress-function, in order that we may ascertain whether and under what conditions these are single-valued. This is particularly necessary in our case, as one of the co-ordinates, β, is itself many-valued.

Adding and subtracting the first two equations (4), and leaving the third as it stands, substituting for the stresses in terms of the stress-function, and for the strains in terms of the displacements, we obtain the following three equations:—

$$\frac{\partial}{\partial\alpha}\left\{\frac{\partial\chi}{\partial\alpha} - 2(\lambda+\mu)\frac{u}{h}\right\} + \frac{\partial}{\partial\beta}\left\{\frac{\partial\chi}{\partial\beta} - 2(\lambda+\mu)\frac{v}{h}\right\} = 0, \quad . \quad . \quad . \quad . \quad . \quad (9)$$

$$\frac{\partial}{\partial\alpha}\left\{h^2\frac{\partial\chi}{\partial\alpha} + 2\mu h u\right\} - \frac{\partial}{\partial\beta}\left\{h^2\frac{\partial\chi}{\partial\beta} + 2\mu h v\right\} = 0, \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

$$\frac{\partial}{\partial\alpha}\left\{h^2\frac{\partial\chi}{\partial\beta} + 2\mu h v\right\} + \frac{\partial}{\partial\beta}\left\{h^2\frac{\partial\chi}{\partial\alpha} + 2\mu h u\right\} = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

From the last two of these it appears that we may define a new function P such that

$$\frac{\partial P}{\partial\alpha} = h^2\frac{\partial\chi}{\partial\beta} + 2\mu h v, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12)$$

$$\frac{\partial P}{\partial\beta} = h^2\frac{\partial\chi}{\partial\alpha} + 2\mu h u, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (13)$$

$$\nabla^2 P = 0,$$

and we have still to satisfy (9). Substituting for u, v in terms of P we have

$$h^2\frac{\partial}{\partial\alpha}\left(\frac{1}{h^2}\frac{\partial P}{\partial\beta}\right)+h^2\frac{\partial}{\partial\beta}\left(\frac{1}{h^2}\frac{\partial P}{\partial\alpha}\right)=\frac{\lambda+2\mu}{\lambda+\mu}\nabla^2\chi,$$

which may be re-arranged thus—

$$\frac{\partial}{\partial\alpha}\left\{\frac{\lambda+2\mu}{\lambda+\mu}\frac{\partial\chi}{\partial\alpha}-\frac{1}{h^2}\frac{\partial P}{\partial\beta}\right\}+\frac{\partial}{\partial\beta}\left\{\frac{\lambda+2\mu}{\lambda+\mu}\frac{\partial\chi}{\partial\beta}-\frac{1}{h^2}\frac{\partial P}{\partial\alpha}\right\}=0.$$

It follows that a function Q exists such that

$$h^2\frac{\partial Q}{\partial\alpha}=\frac{\partial P}{\partial\alpha}-\frac{\lambda+2\mu}{\lambda+\mu}h^2\frac{\partial\chi}{\partial\beta}, \quad \ldots\ldots \quad (14)$$

$$h^2\frac{\partial Q}{\partial\beta}=-\frac{\partial P}{\partial\beta}+\frac{\lambda+2\mu}{\lambda+\mu}h^2\frac{\partial\chi}{\partial\alpha}. \quad \ldots\ldots \quad (15)$$

Eliminating P by differentiating with regard to β and α respectively, and adding, we have

$$\frac{\partial^2(hQ)}{\partial\alpha\,\partial\beta}-Q\frac{\partial^2 h}{\partial\alpha\,\partial\beta}=\frac{\lambda+2\mu}{2(\lambda+\mu)}\left\{\frac{1}{h}\frac{\partial}{\partial\alpha}\left(h^2\frac{\partial\chi}{\partial\alpha}\right)-\frac{1}{h}\frac{\partial}{\partial\beta}\left(h^2\frac{\partial\chi}{\partial\beta}\right)\right\},$$

which becomes in our co-ordinates

$$\frac{\partial^2(hQ)}{\partial\alpha\,\partial\beta}=\frac{\lambda+2\mu}{2(\lambda+\mu)}\left\{\frac{\partial^2(h\chi)}{\partial\alpha^2}-\frac{\partial^2(h\chi)}{\partial\beta^2}-h\chi\right\}. \quad \ldots\ldots \quad (16)$$

There is, however, a further condition to be satisfied by Q corresponding to the condition $\nabla^2P=0$. Differentiating (14) and (15) with regard to α, β respectively, and subtracting we have

$$\frac{\partial}{\partial\alpha}\left(h^2\frac{\partial Q}{\partial\alpha}\right)-\frac{\partial}{\partial\beta}\left(h^2\frac{\partial Q}{\partial\beta}\right)=-\frac{\lambda+2\mu}{\lambda+\mu}\left\{\frac{\partial}{\partial\alpha}\left(h^2\frac{\partial\chi}{\partial\beta}\right)+\frac{\partial}{\partial\beta}\left(h^2\frac{\partial\chi}{\partial\alpha}\right)\right\},$$

or in our co-ordinates

$$\frac{\partial^2}{\partial\alpha^2}(hQ)-\frac{\partial^2}{\partial\beta^2}(hQ)-hQ=-\frac{2(\lambda+2\mu)}{\lambda+\mu}\frac{\partial^2(hQ)}{\partial\alpha\,\partial\beta}. \quad \ldots\ldots \quad (17)$$

These two equations connecting Q and χ are consistent, for, if we eliminate Q by appropriate differential operators, we have

$$\left\{\frac{\partial^2}{\partial\alpha^2}-\frac{\partial^2}{\partial\beta^2}-1\right\}^2(h\chi)=-4\frac{\partial^4(h\chi)}{\partial\alpha^2\,\partial\beta^2},$$

which is readily seen to be identical with the condition $\nabla^4\chi=0$, as given in (8). It is obvious that hQ satisfies the same differential equation, and hence it also is a solution of $\nabla^4Q=0$.

We have therefore from (16)

$$hQ = \frac{\lambda+2\mu}{2(\lambda+\mu)}\iint\left\{\frac{\partial^2(h\chi)}{\partial\alpha^2}-\frac{\partial^2(h\chi)}{\partial\beta^2}-h\chi\right\}d\alpha\,d\beta, \qquad (18)$$

and from (12), (13) and (14), (15)

$$2\mu u = \frac{\mu}{\lambda+\mu}h\frac{\partial\chi}{\partial\alpha}-h\frac{\partial Q}{\partial\beta}, \qquad (19)$$

$$2\mu v = \frac{\mu}{\lambda+\mu}h\frac{\partial\chi}{\partial\beta}+h\frac{\partial Q}{\partial\alpha}. \qquad (20)$$

It is readily seen that these equations determine u and v apart possibly from rigid body displacements, for, although owing to the double integration an arbitrary function of α and an arbitrary function of β will appear in hQ, these will be determined by (17), except for functions of α or β, which make its left-hand side vanish identically. The only possible arbitrary terms in hQ are therefore given by $hQ = a\mathrm{A}(\cosh\alpha+\cos\beta)+\mathrm{B}(\cosh\alpha-\cos\beta)+\mathrm{C}a\sinh\alpha+\mathrm{D}a\sin\beta$, or

$$Q = \mathrm{A}r^2+a\mathrm{B}+\mathrm{C}y+\mathrm{D}x$$

where r is the distance from the origin. These give rise to terms in u, v corresponding to motions of pure translation and rigid body rotation about the origin.

§ 3. THE STRESS-FUNCTION.

Turning now to the consideration of the possible forms for the stress-function in these co-ordinates, we note that the differential equation (8) can readily be solved by the ordinary method, and that its general solution is

$$h\chi = e^{\alpha}\phi_1(\alpha+i\beta)+e^{-\alpha}\phi_2(\alpha+i\beta)+e^{\alpha}\phi_3(\alpha-i\beta)+e^{-\alpha}\phi_4(\alpha-i\beta).$$

If we seek a solution of the type $h\chi = f(\alpha)\cos n\beta$ or $f(\alpha)\sin n\beta$, (8) shows that the differential equation for $f(\alpha)$ is

$$\left(\frac{d^4}{d\alpha^4}-2(n^2+1)\frac{d^2}{d\alpha^2}+n^4-2n^2+1\right)f(\alpha) = 0,$$

the solution of which is

$$f(\alpha) = \mathrm{A}_n\cosh(n+1)\alpha+\mathrm{B}_n\cosh(n-1)\alpha+\mathrm{C}_n\sinh(n+1)\alpha+\mathrm{D}_n\sinh(n-1)\alpha,$$

unless $n = 0$ or 1. In the latter case we have

$$f(\alpha) = \mathrm{A}_1\cosh 2\alpha+\mathrm{B}_1+\mathrm{C}_1\sinh 2\alpha+\mathrm{D}_1\alpha,$$

and when $n = 0$

$$f(\alpha) = \mathrm{A}_0\cosh\alpha+\mathrm{B}_0\alpha\cosh\alpha+\mathrm{C}_0\sinh\alpha+\mathrm{D}_0\alpha\sinh\alpha.$$

If we now seek solutions for which $h\chi$ is a multiple of $\sinh n\alpha$ or $\cosh n\alpha$, we find the following solutions which are not included above—

$$h\chi = (\mathrm{E}\cos\beta+\mathrm{F}\sin\beta+\mathrm{G}\cosh\alpha+\mathrm{H}\sinh\alpha)\,\beta.$$

Since any constant multiples of x, y and any constant may be added to χ without affecting the stresses, it follows from (2) that any multiples of

$$\sinh\alpha,\ \sin\beta \ \text{or}\ \cosh\alpha-\cos\beta$$

may be added to $h\chi$. This allows us to take the coefficients of $\cosh\alpha$, $\sinh\alpha$ and $\sin\beta$ as zero. We have then the following general expression for $h\chi$:—

$$\begin{aligned} h\chi = {} & (\mathrm{E}\cos\beta+\mathrm{F}\sin\beta+\mathrm{G}\cosh\alpha+\mathrm{H}\sinh\alpha)\,\beta \\ & +(\mathrm{B}_0\cosh\alpha+\mathrm{D}_0\sinh\alpha)\,\alpha \\ & +(\mathrm{A}_1\cosh 2\alpha+\mathrm{B}_1+\mathrm{C}_1\sinh 2\alpha+\mathrm{D}_1\alpha)\cos\beta \\ & +(\mathrm{A}'_1\cosh 2\alpha+\mathrm{C}'_1\sinh 2\alpha+\mathrm{D}'_1\alpha)\sin\beta \\ & +\sum_{n=2}^{\infty}\left\{\begin{aligned} & [\mathrm{A}_n\cosh(n+1)\,\alpha+\mathrm{B}_n\cosh(n-1)\,\alpha+\mathrm{C}_n\sinh(n+1)\,\alpha \\ & \qquad\qquad +\mathrm{D}_n\sinh(n-1)\,\alpha]\cos n\beta \\ & +[\mathrm{A}'\cosh(n+1)\,\alpha+\mathrm{B}'_n\cosh(n-1)\,\alpha+\mathrm{C}'_n\sinh(n+1)\,\alpha \\ & \qquad\qquad +\mathrm{D}'_n\sinh(n-1)\,\alpha]\sin n\beta. \end{aligned}\right\} \end{aligned} \qquad . \quad . \quad . \quad (21)$$

We have now to determine whether the displacements corresponding to this stress-function are single valued or not. The function $(h\mathrm{Q})$ is easily obtained by simple integration from (16), and the arbitrary functions thus appearing can be determined by the aid of (17). We have

$$\begin{aligned} -\frac{\lambda+\mu}{\lambda+2\mu}(h\mathrm{Q}) = {} & (\mathrm{E}\cos\beta+\mathrm{F}\sin\beta+\mathrm{G}\cosh\alpha+\mathrm{H}\sinh\alpha)\,\alpha \\ & -(\mathrm{B}_0\cosh\alpha+\mathrm{D}_0\sinh\alpha)\,\beta \\ & -(\mathrm{A}_1\sinh 2\alpha+\mathrm{C}_1\cosh 2\alpha+\mathrm{D}'_1\beta)\sin\beta \\ & +(\mathrm{A}'_1\sinh 2\alpha+\mathrm{C}'_1\cosh 2\alpha-\mathrm{D}_1\beta)\cos\beta \\ & +\sum_{n=2}^{\infty}\left\{\begin{aligned} & [\mathrm{A}'_n\sinh(n+1)\,\alpha+\mathrm{B}'_n\sinh(n-1)\,\alpha+\mathrm{C}'_n\cosh(n+1)\,\alpha \\ & \qquad\qquad +\mathrm{D}'_n\cosh(n-1)\,\alpha]\cos n\beta \\ & -[\mathrm{A}_n\sinh(n+1)\,\alpha+\mathrm{B}_n\sinh(n-1)\,\alpha+\mathrm{C}_n\cosh(n+1)\,\alpha \\ & \qquad\qquad +\mathrm{D}_n\cosh(n-1)\,\alpha]\sin n\beta. \end{aligned}\right\} \end{aligned} \qquad . \quad . \quad . \quad (22)$$

It is clear, from the general expressions for $h\chi$ and $h\mathrm{Q}$, that the only terms which can possibly give rise to many-valued displacements are

$$\begin{aligned} h\chi = {} & (\mathrm{E}\cos\beta+\mathrm{F}\sin\beta+\mathrm{G}\cosh\alpha+\mathrm{H}\sinh\alpha)\,\beta \\ & +(\mathrm{B}_0\cosh\alpha+\mathrm{D}_0\sinh\alpha+\mathrm{D}_1\cos\beta+\mathrm{D}'_1\sin\beta)\,\alpha, \end{aligned}$$

and the corresponding terms in hQ

$$-\frac{\lambda+\mu}{\lambda+2\mu}hQ = (E\cos\beta+F\sin\beta+G\cosh\alpha+H\sinh\alpha)\,\alpha$$
$$-(B_0\cosh\alpha+D_0\sinh\alpha+D_1\cos\beta+D'_1\sin\beta)\,\beta.$$

From (19) and (20) we may now find the corresponding displacements u, v. Each of these is found to contain a multiple of the many-valued co-ordinate β. Equating the coefficients of these terms to zero we have the following relations :—

$$\left.\begin{aligned} E+G &= 0, & B_0+D_1 &= 0, \\ \mu F-(\lambda+2\mu)\,D_0 &= 0, & \mu H+(\lambda+2\mu)\,D'_1 &= 0. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad (23)$$

We shall now show that these early terms correspond to the resultant of the forces and couple applied over the boundaries. For this purpose we shall require the following elementary forms of the stress-function :—

(1) For an isolated force X applied at the origin in the direction of the x-axis

$$\chi = -(2\pi)^{-1}X\,(y\theta-\nu x\log r)$$

where r, θ as usual denote polar co-ordinates and $\nu = \mu/(\lambda+2\mu)$.

(2) For an isolated force Y applied at the origin in the direction of the y-axis

$$\chi = (2\pi)^{-1}Y\,(x\theta+\nu y\log r).$$

(3) For a point couple of moment L applied at the origin in a positive sense

$$\chi = -(2\pi)^{-1}L\theta.$$

(4) For a centre of pressure radiating uniformly from the origin

$$\chi = \log r,$$

Inserting the relations (23) necessary to ensure single-valued displacements our early terms become

$$h\chi = G\,(\cosh\alpha-\cos\beta)\,\beta+\beta_0\,(\cosh\alpha-\cos\beta)\,\alpha$$
$$+F\,(\beta\sin\beta+\nu\alpha\sinh\alpha)+H\,(\beta\sinh\alpha-\nu\alpha\sin\beta)$$

or

$$\chi = aG\beta+aB_0\alpha+F\,(x\beta+\nu y\alpha)+H\,(y\beta-\nu x\alpha). \quad . \quad . \quad . \quad . \quad . \quad (24)$$

Now $aG\beta = aG\,(\theta_1-\theta_2)$ and hence this term represents a couple of moment $2\pi aG$ applied at $\alpha = \infty$ and an equal and opposite couple applied at $\alpha = -\infty$. The term $a\beta_0\alpha$ represents two equal and opposite centres of radial pressure at these same points,

We also have

$$\mathrm{H}(y\beta-\nu x\alpha) = \mathrm{H}\{(y+a)\,\theta_1-\nu x\log r_1\}-\mathrm{H}\{(y-a)\,\theta_2-\nu x\log r_2\}$$
$$-a\mathrm{H}(\theta_1+\theta_2).$$

This corresponds to a force $2\pi\mathrm{H}$ applied at $\alpha = +\infty$ parallel to the x-axis and an equal and opposite force applied at $\alpha = -\infty$ (thus forming a couple of moment $4\pi a\mathrm{H}$) and point couples each of moment $2\pi a\mathrm{H}$ applied at these same points (see fig. 2).

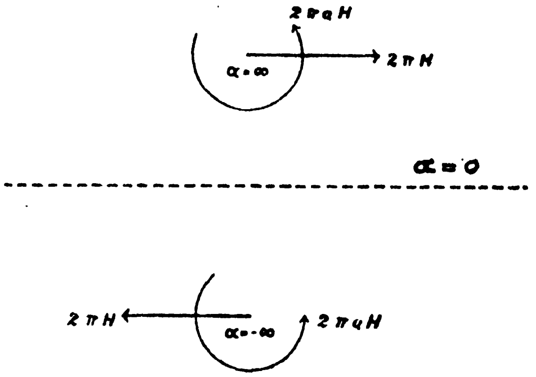


Fig. 2.

Finally

$$\mathrm{F}\; x\beta+\nu y\alpha) = \mathrm{F}\{x\theta_1+\nu(y+a)\log r_1\}$$
$$-\mathrm{F}\{x\theta_2+\nu(y-a)\log r_2\}$$
$$-a\nu\mathrm{F}\log r_1 r_2.$$

This corresponds to forces each equal to $2\pi\mathrm{F}$, acting at the points $\alpha = \pm\infty$ and each directed towards the origin, together with two equal like centres of uniform pressure at the same points. This brings to light a new solution corresponding to the last term.

Expressed in our co-ordinates we have

$$\log r_1 r_2 = 2\log(2a)-2\log(\cosh\alpha-\cos\beta),$$

and the corresponding form of $h\chi$ is, apart from constants,

$$h\chi = (\cosh\alpha-\cos\beta)\log(\cosh\alpha-\cos\beta).$$

It is easily seen that this can be expanded in a Fourier series which is included in our general expression for $h\chi$, but that the expansion is different on opposite sides of the line $\alpha = 0$. For this reason we shall find it convenient to include a term of this form whenever the region under consideration includes parts above and below the axis of x, *i.e.*, when it is bounded by two circles neither of which encloses the other.

It will be noted that, taken together, the early terms allow for the most general resultant forces acting over the two circular boundaries enclosing the two points $\alpha = +\infty$, $\alpha = -\infty$, subject to the condition that the forces acting over the two boundaries considered together form a system in equilibrium. If it is desired to investigate problems for which this condition is not satisfied we can readily obtain the necessary additional solutions. They will be

$$\left.\begin{aligned}\chi &= (y+a)\,\theta_1-\nu x\log r_1\\ \chi &= x\theta_1+\nu(y+\alpha)\log r_1\\ \chi &= \theta_1\end{aligned}\right\}\quad\ldots\ldots\ldots\quad(25)$$

corresponding to forces and couple applied at $\alpha = -\infty$, and similar terms in θ_2, $\log r_2$ corresponding to forces and couple applied at $\alpha = +\infty$. The corresponding forms of $h\chi$ can be expanded in series which are included in our general form, but here again the expansions are different on opposite sides of $\alpha = 0$ and diverge for $\alpha = 0$, $\beta = 0$ together, *i.e.*, at infinity. This divergence corresponds to the obvious fact that forces or couples must be applied at infinity to maintain equilibrium.

Owing to difficulties of this kind we shall find it convenient to insert the appropriate terms corresponding to the resultant force and couple over a boundary and to investigate the stress-function corresponding to the remaining applied forces which will be in statical equilibrium for each boundary.

Let us write for brevity

$$\left.\begin{aligned}\phi_n(\alpha) &= \mathrm{A}_n \cosh(n+1)\alpha + \mathrm{B}_n \cosh(n-1)\alpha + \mathrm{C}_n \sinh(n+1)\alpha + \mathrm{D}_n \sinh(n-1)\alpha \\ \psi_n(\alpha) &= \mathrm{A}'_n \cosh(n+1)\alpha + \mathrm{B}'_n \cosh(n-1)\alpha + \mathrm{C}'_n \sinh(n+1)\alpha + \mathrm{D}_n' \sinh(n-1)\alpha,\end{aligned}\right\} \qquad (26)$$

if $n \geqq 2$ and

$$\left.\begin{aligned}\phi_1(\alpha) &= \mathrm{A}_1 \cosh 2\alpha + \mathrm{B}_1 + \mathrm{C}_1 \sinh 2\alpha \\ \psi_1(\alpha) &= \mathrm{A}'_1 \cosh 2\alpha + \mathrm{C}'_1 \sinh 2\alpha.\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (27)$$

Setting aside the terms corresponding to the resultant forces and couples over the separate boundaries we have

$$\begin{aligned}h\chi = {} & \{\mathrm{B}_0\alpha + \mathrm{K} \log(\cosh\alpha - \cos\beta)\}(\cosh\alpha - \cos\beta) \\ & + \sum_{n=1}^{\infty} \{\phi_n(\alpha)\cos n\beta + \psi_n(\alpha)\sin n\beta\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (28)\end{aligned}$$

where the term in K may be omitted when the region considered lies entirely on one side of the line

§ 4. BOUNDARY CONDITIONS.

Let us consider a plate bounded by two curves $\alpha = \alpha_1, \alpha_2$. We may suppose $\alpha_1 > \alpha_2$ and $\alpha_1 > 0$. Then, if $\alpha_2 > 0$ we have a finite plate bounded internally and externally by circles which are not concentric, if $\alpha_2 < 0$ we have an infinite plate containing two circular holes, and by suitably choosing the values of a, α_1, α_2 we can make the circular boundaries in either case of any desired radius and centre distance. In particular if $\alpha_2 = 0$, we have a semi-infinite plate bounded by a straight edge and containing a circular hole. Suppose that such a plate is in equilibrium under given normal and tangential forces applied over the boundaries $\alpha = \alpha_1, \alpha_2$, so that we are given over $\alpha = \alpha_1$,

$$\left.\begin{aligned}a\widehat{\alpha\beta} &= a_0 + \sum_1^{\infty}(a_n \cos n\beta + b_n \sin n\beta), \\ a\widehat{\alpha\alpha} &= c_0 + \sum_1^{\infty}(c_n \cos n\beta + d_n \sin n\beta),\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (29)$$

while over $\alpha = \alpha_2$ we have similar expansions in which a_0, a_n, b_n, c_0, c_n, d_n are replaced by a'_0, a'_n, b'_n, c'_0, c'_n, d'_n.

If the tractions applied over the circle $\alpha = \alpha_1$ are statically equivalent to forces X, Y at its centre, and a couple of moment L, then

$$X = \int_0^{2\pi}\left\{\widehat{\alpha\alpha}\frac{\partial x}{\partial \alpha} - \widehat{\alpha\beta}\frac{\partial y}{\partial \alpha}\right\}d\beta,$$

$$Y = \int_0^{2\pi}\left\{\widehat{\alpha\alpha}\frac{\partial y}{\partial \alpha} + \widehat{\alpha\beta}\frac{\partial x}{\partial \alpha}\right\}d\beta,$$

$$L = -\frac{a^2}{\sinh \alpha_1}\int_0^{2\pi}\frac{\widehat{\alpha\beta}\,d\beta}{\cosh \alpha - \cos \beta}.$$

The coefficients of $\widehat{\alpha\alpha}$, $\widehat{\beta\beta}$ can readily be expanded in Fourier series. We have, in fact, since $\alpha_1 > 0$,

$$\frac{\partial x}{\partial \alpha} = -a\frac{\sinh \alpha \sin \beta}{(\cosh \alpha - \cos \beta)^2} = -2a\sum_1^\infty ne^{-n\alpha_1}\sin n\beta,$$

$$\frac{\partial y}{\partial \alpha} = -a\frac{(\cosh \alpha \cos \beta - 1)}{(\cosh \alpha - \cos \beta)^2} = -2a\sum_1^\infty ne^{-n\alpha_1}\cos n\beta,$$

and

$$\sinh \alpha_1(\cosh \alpha_1 - \cos \beta)^{-1} = 1 + 2\sum_1^\infty e^{-n\alpha_1}\cos n\beta.$$

Substituting these and the expansions for $\widehat{\alpha\alpha}$, $\widehat{\alpha\beta}$ in the expressions for X, Y, L, and integrating, we have

$$X = 2\pi\sum_1^\infty n(a_n - d_n)e^{-n\alpha_1},$$

$$Y = -2\pi\sum_1^\infty n(b_n + c_n)e^{-n\alpha_1},$$

$$L = -2\pi a \operatorname{cosech}^2\alpha_1 \sum_1^\infty a_n e^{-n\alpha_1}.$$

The corresponding components of the resultant of the forces applied over $\alpha = \alpha_2$ can be obtained in a similar way. We must, however, remember that in this case the forces act from that side of the boundary for which $\alpha < \alpha_2$, whereas in the case of the first boundary they acted from the side for which $\alpha < \alpha_1$. We obtain, if $\alpha_2 > 0$,

$$X' = -2\pi\sum_1^\infty n(a'_n - d'_n)e^{-n\alpha_2},$$

$$Y' = 2\pi\sum_1^\infty n(b'_n + c'_n)e^{-n\alpha_2},$$

$$L' = 2\pi a \operatorname{cosech}^2\alpha_2 \sum_1^\infty a'_n e^{-n\alpha_2}.$$

If $\alpha_2 < 0$ there are some differences of sign owing to the different Fourier expansions for the direction cosines. We have

$$X' = -2\pi \sum_1^\infty n\,(a'_n + d'_n)\, e^{n\alpha_2},$$

$$Y' = -2\pi \sum_1^\infty n\,(b'_n - c'_n)\, e^{n\alpha_2},$$

$$L' = -2\pi a \operatorname{cosech}^2\alpha_2 \sum_1^\infty a'_n e^{n\alpha_2}.$$

Hence, if the forces acting on each boundary are statically in equilibrium, we have

$$\left.\begin{aligned} &\sum_1^\infty n\,(a_n - d_n)\, e^{-n\alpha_1} = 0, \qquad &&\sum_1^\infty n\,(b_n + c_n)\, e^{-n\alpha_1} = 0, \\ &\sum_1^\infty a_n e^{-n\alpha_1} = 0, &&\sum_1^\infty a'_n e^{\pm n\alpha_2} = 0. \end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (30)$$

with, if $\alpha_2 > 0$,

$$\sum_1^\infty n\,(a'_n - d'_n)\, e^{-n\alpha_2} = 0, \qquad \sum_1^\infty n\,(b'_n + c'_n)\, e^{-n\alpha_2} = 0, \quad . \quad . \quad . \quad . \quad (31)$$

or, if $\alpha_2 < 0$,

$$\sum_1^\infty n\,(a'_n + d'_n)\, e^{n\alpha_2} = 0, \qquad \sum_1^\infty n\,(b'_v - c'_n)\, e^{n\alpha_2} = 0. \quad . \quad . \quad . \quad . \quad . \quad (32)$$

We will now show that it is possible to determine a stress-function of the form (28) which gives the appropriate stresses over $\alpha = \alpha_1, \alpha_2$, and which gives no stress at infinity if the region considered extends so far.

By the aid of (6) we can calculate the stresses corresponding to the stress-function (28). We obtain

$$\begin{aligned} 2a\widehat{\alpha\alpha} = {} & \mathrm{K}\,(1 - 2\cosh^2\alpha) - 2\mathrm{B}_0 \sinh\alpha \cosh\alpha + 2\phi_1(\alpha) \\ & + 2\,(\mathrm{K}\cosh\alpha + \mathrm{B}_0 \sinh\alpha)\cos\beta - \mathrm{K}\cos 2\beta \\ & + \sum_{n=1}^\infty \left\{ \begin{aligned} & [(n+1)(n+2)\,\phi_{n+1}(\alpha) - 2\cosh\alpha\,(n^2-1)\,\phi_n(\alpha) \\ & \qquad\qquad + (n-1)(n-2)\,\phi_{n-1}(\alpha)]\cos n\beta \\ & + [(n+1)(n+2)\,\psi_{n+1}(\alpha) - 2\cosh\alpha\,(n^2-1)\,\psi_n(\alpha) \\ & \qquad\qquad + (n-1)(n-2)\,\psi_{n-1}(\alpha)]\sin n\beta \\ & - 2\sinh\alpha\,[\phi'_n(\alpha)\cos n\beta + \psi'_n(\alpha)\sin n\beta]. \end{aligned} \right\} \end{aligned}$$

and

$$\begin{aligned} 2a\widehat{\alpha\beta} = {} & \psi'_1(\alpha) - 2\,(\mathrm{K}\sinh\alpha + \mathrm{B}_0\cosh\alpha)\sin\beta + \mathrm{B}_0 \sin 2\beta \\ & + \sum_{n=1}^\infty \left\{ \begin{aligned} & [(n+1)\,\psi'_{n+1}(\alpha) - 2\cosh\alpha\, n\psi'_n(\alpha) + (n-1)\,\psi'_{n-1}(\alpha)]\cos n\beta, \\ & - [(n+1)\,\phi'_{n+1}(\alpha) - 2\cosh\alpha\, n\phi'_n(\alpha) + (n-1)\,\phi'_{n-1}(\alpha)]\sin n\beta. \end{aligned} \right\} \end{aligned}$$

Identifying these with (29), we have the following relations from which to obtain the coefficients :—

$$\left.\begin{aligned}
&\phi_1(\alpha_1) = 2c_0+2\mathrm{B}_0 \sinh\alpha_1 \cos\alpha_1+2\mathrm{K}\cosh^2\alpha_1-\mathrm{K}\\
&1.2.3.\phi_2(\alpha_1) = 2c_1-2(\mathrm{K}\cosh\alpha_1+\mathrm{B}_0\sinh\alpha_1)+2\sinh\alpha_1\,\phi'_1(\alpha_1)\\
&2.3.4.\phi_3(\alpha_1)-2\cosh\alpha_1\,.\,1.2.3.\phi_2(\alpha_1) = 4c_2+2\mathrm{K}+4\sinh\alpha_1\,\phi'_2(\alpha_1)\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&n(n+1)(n+2)\phi_{n+1}(\alpha_1)-2\cosh\alpha_1(n-1)(n)(n+1)\phi_n(\alpha_1)+(n-2)(n-1)(n)\phi_{n-1}(\alpha_1)\\
&\qquad\qquad = 2nc_n+2n\sinh\alpha_1\,\phi'_n(\alpha_1) \qquad\qquad (n\geqq 3)
\end{aligned}\right\}\quad(33)$$

$$\left.\begin{aligned}
&n(n+1)(n+2)\psi_{n+1}(\alpha_1)-2\cosh\alpha_1(n-1)(n)(n+1)\psi_n(\alpha_1)+(n-2)(n-1)(n)\psi_{n-1}(\alpha_1)\\
&\qquad\qquad = 2nd_n+2n\sinh\alpha_1\,\psi'_n(\alpha_1) \qquad\qquad (n\geqq 1)
\end{aligned}\right\}\quad . \; . \; . \quad(34)$$

$$\left.\begin{aligned}
&\psi'_1(\alpha_1) = 2a_0\\
&(n+1)\psi'_{n+1}(\alpha_1)-2\cosh\alpha_1\,n\psi'_n(\alpha_1)+(n-1)\psi'_{n-1}(\alpha_1) = 2a_n \qquad (n\geqq 1)
\end{aligned}\right\}\quad . \quad(35)$$

$$\left.\begin{aligned}
&2\phi'_2(\alpha_1)-2\cosh\alpha_1\,\phi'_1(\alpha_1) = -2b_1-2(\mathrm{K}\sinh\alpha_1+\mathrm{B}_0\cosh\alpha_1)\\
&3\phi'_3(\alpha_1)-4\cosh\alpha_1\,\phi'_2(\alpha_1)+\phi'_1(\alpha_1) = -2b_2+\mathrm{B}_0\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&(n+1)\phi'_{n+1}(\alpha_1)-2\cosh\alpha_1\,n\phi'_n(\alpha_1)+(n-1)\phi'_{n-1}(\alpha_1) = -2b_n \qquad (n\geqq 3)
\end{aligned}\right\}\quad . \quad(36)$$

Writing out equations (35), multiplying by $e^{-n\alpha_1}$ and adding, we have

$$(n+1)\,\psi'_{n+1}(\alpha_1)\,e^{-n\alpha_1}-n\psi'_n(\alpha_1)\,e^{-(n+1)\alpha_1} = 2\sum_{p=0}^{n} a_p e^{-p\alpha_1}$$

or

$$(n+1)\,\psi'_{n+1}(\alpha_1)-n\psi'_n(\alpha_1)\,e^{-\alpha_1} = 2e^{n\alpha_1}\sum_{p=0}^{n} a_p e^{-p\alpha_1}. \quad . \; . \; . \quad (37)$$

Now, in virtue of (30), we may write the right-hand side of this

$$-2e^{n\alpha_1}\sum_{p=n+1}^{\infty} a_p e^{-p\alpha_1} = -2\sum_{r=1}^{\infty} a_{n+r}e^{-r\alpha_1},$$

and since $\Sigma a_n \cos n\beta$ is supposed convergent this tends to zero as n increases. Hence from (37) we see that the limit of $\psi'_{n+1}(\alpha_1)/\psi'_n(\alpha_1)$ as n increases is $e^{-\alpha_1}$, and hence the functions $\psi'_n(\alpha_1)$ are finite for all values of n and tend to zero as n increases, if the resultant couple acting on $\alpha = \alpha_1$ vanishes.

Multiplying (37) by $e^{n\alpha_1}$ and adding, we have

$$n\psi'_n(\alpha_1)\,e^{(n-1)\alpha_1} = 2\sum_{q=0}^{n-1}\sum_{p=0}^{q} a_p e^{(2q-p)\alpha_1} = 2\sum_{p=0}^{n-1}\sum_{q=p}^{n-1} a_p e^{(2q-p)\alpha_1},$$

which, on effecting the summation with regard to q, leads at once to

$$n\psi'_n(\alpha_1) = 2\,\mathrm{cosech}\,\alpha_1 \sum_{p=0}^{n-1} a_p \sinh(n-p)\alpha_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (38)$$

for $n \geqq 1$.

Treating (34) in a similar way we have

$$n(n+1)(n+2)\psi_{n+1}(\alpha_1)-(n-1)(n)(n+1)\psi_n(\alpha_1)e^{-\alpha_1}$$
$$= 2e^{n\alpha_1}\sum_{p=1}^{n} p\left[d_p+\sinh\alpha_1\psi'_p(\alpha_1)\right]e^{-p\alpha_1}. \quad . \quad . \qquad (39)$$

We can readily show from (35) that

$$\sinh\alpha_1 \sum_{p=1}^{n} pe^{-p\alpha_1}\psi'_p(\alpha_1) = \tfrac{1}{2}n(n+1)\{\psi'_{n+1}(\alpha_1)-e^{-\alpha}\psi'_n(\alpha_1)\}e^{-n\alpha_1}-\sum_{p=1}^{n} pa_pe^{-p\alpha_1},$$

and hence (39) may be written

$$n(n+1)(n+2)\psi_{n+1}(\alpha_1)-(n-1)(n)(n+1)\psi_n(\alpha_1)e^{-\alpha_1} = n(n+1)\left[\psi'_{n+1}(\alpha_1)+e^{-\alpha_1}\psi'_n(\alpha_1)\right]$$
$$+2e^{n\alpha_1}\sum_{p=1}^{n} p(d_p-a_p)e^{-p\alpha_1}.$$

As in the case of $\psi'_n(\alpha_1)$, we can show that the right-hand side tends to zero as n increases if conditions (30) are fulfilled. Hence $\psi_n(\alpha_1)$ is finite for all values of n and tends to zero as n increases, and we have

$$(n-1)(n)(n+1)\psi_n(\alpha_1)e^{(n-1)\alpha_1} = 2\sum_{q=1}^{n-1}\sum_{p=1}^{q} p(d_p-a_p)e^{(2q-p)\alpha_1}$$
$$+\sum_{q=1}^{n-1} q(q+1)e^{q\alpha_1}\left[\psi'_{q+1}(\alpha_1)-e^{-\alpha_1}\psi'_q(\alpha_1)\right],$$

which, on reduction, leads to

$$n(n^2-1)\psi_n(\alpha_1) = 2\,\mathrm{cosech}\,\alpha_1\sum_{p=0}^{n-1}\{(n-p)a_p\cosh(n-p)\alpha_1$$
$$+(pd_p-a_p\coth\alpha_1)\sinh(n-p)\alpha_1\} \quad . \quad . \qquad (40)$$

for $n \geqq 2$. Equations (34) do not determine $\psi_1(\alpha_1)$.

From (36) we have

$$2\phi'_2(\alpha_1)e^{-\alpha_1}-\phi'_1(\alpha_1)e^{-2\alpha_1} = \phi'_1(\alpha_1)-2e^{-\alpha_1}(\mathrm{K}\sinh\alpha_1+\mathrm{B}_0\cosh\alpha_1)-2b_1e^{-\alpha_1}, \quad . \qquad (41)$$

and if $n \geqq 2$

$$(n+1)\phi'_{n+1}(\alpha_1)e^{-n\alpha_1}-n\phi'_n(\alpha_1)e^{-(n+1)\alpha_1} = \phi'_1(\alpha_1)-2\mathrm{K}e^{-\alpha_1}\sinh\alpha_1-\mathrm{B}_0$$
$$-2\sum_{p=1}^{n} b_pe^{-p\alpha_1}, \quad . \quad , \quad . \quad . \quad . \quad . \qquad (42)$$

and hence, if the sequence $\phi_n(\alpha_1)$ is to converge for large values of n, we must have

$$\phi'_1(\alpha_1) = \mathrm{B}_0+2\mathrm{K}e^{-\alpha_1}\sinh\alpha_1+2\sum_{p=1}^{\infty} b_pe^{-p\alpha_1}. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (43)$$

From (41) and (42) we have, if $n \geqq 2$,

$$n\phi'_n(\alpha_1)\, e^{(n-1)\alpha_1} = \phi'_1(\alpha_1) - \mathrm{B}_0 + \sum_{q=1}^{n-1} (\phi'_1(\alpha_1) - 2\mathrm{K}e^{-\alpha_1} \sinh \alpha_1 - \mathrm{B}_0)\, e^{2q\alpha_1}$$
$$-2 \sum_{q=1}^{n-1} \sum_{p=1}^{q} b_p e^{(2q-p)\alpha_1},$$

from which we obtain

$$n \sinh \alpha_1\, \phi'_n(\alpha_1) = (\phi'_1(\alpha_1) - \mathrm{B}_0) \sinh n\alpha_1 - 2\mathrm{K} \sinh (n-1)\alpha_1 \sinh \alpha_1$$
$$-2 \sum_{v=1}^{n-1} b_p \sinh (n-p)\alpha_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (44)$$

for $n \geqq 2$, while $\phi'_1(\alpha_1)$ is given by (43).

Finally we have from (33), omitting the first equation of the series, if $n \geqq 2$,

$$n(n+1)(n+2)\, e^{-n\alpha_1} \phi_{n+1}(\alpha_1) - (n-1)(n)(n+1)\, e^{-(n+1)\alpha_1} \phi_n(\alpha_1)$$
$$= -2(\mathrm{B}_0 + \mathrm{K})\, e^{-\alpha_1} \sinh \alpha_1$$
$$+2 \sum_{p=1}^{n} p\, (c_p + \sinh \alpha_1\, \phi'_p(\alpha_1))\, e^{-p\alpha_1}.$$

By the aid of (36) we can reduce the right-hand side to

$$n(n+1)(\phi'_{n+1}(\alpha_1) - e^{-\alpha_1} \phi'_n(\alpha_1))\, e^{-n\alpha_1} + 2 \sum_{p=1}^{n} p\, (c_p + b_p)\, e^{-p\alpha_1},$$

from which it appears that $\phi_{n+1}(\alpha_1)$ is finite and tends to zero as n increases, provided that the resultant of the applied forces over $\alpha = \alpha_1$ is zero. We then obtain for $n \geqq 2$,

$$n(n^2-1) \sinh \alpha_1\, \phi_n(\alpha_1) = (\phi'_1(\alpha_1) - \mathrm{B}_0) \{n \cosh n\alpha_1 - \coth \alpha_1 \sinh n\alpha_1\}$$
$$-\mathrm{K}\{(n-1) \sinh n\alpha_1 - (n+1) \sinh (n-2)\alpha_1\}$$
$$+2 \sum_{p=1}^{n-1} \{(pc_p + b_p \coth \alpha_1) \sinh (n-p)\alpha_1$$
$$-(n-p)\, b_p \cosh (n-p)\alpha_1\}, \quad . \quad . \quad . \quad . \quad (45)$$

while

$$2\phi_1(\alpha_1) = 2c_0 + \mathrm{B}_0 \sinh 2\alpha_1 + \mathrm{K}(2 \cosh^2 \alpha - 1). \quad . \quad . \quad . \quad . \quad . \quad (46)$$

It appears that equations (38), (40), (43), (44), (45) and (46) give the values of $\phi_n(\alpha_1)$, $\psi_n(\alpha_1)$, $\phi'_n(\alpha_1)$, $\psi'_n(\alpha_1)$ for $n \geqq 1$ in terms of B_0, K and the given coefficients a_n, &c., with the exception of $\psi_1(\alpha_1)$.

Now we have only assumed $\alpha_1 > 0$ in order to establish the convergence of these functions, and hence the corresponding functions of α_2 will be given by the same formulæ with a'_n, b'_n, c'_n, d'_n substituted for a_n, b_n, c_n, d_n, provided that the conditions for convergence are satisfied. It may be shown that the new conditions of convergence are identical with (30) and (31), or (30) and (32), according as $\alpha_2 >$ or < 0.

The formula for $\phi_1(\alpha_1)$ given in (43), which is itself a condition of convergence, will, however, be replaced by

$$\phi'_1(\alpha_2) = B_0 + 2Ke^{\alpha_2} \sinh \alpha_2 + 2\sum_{n=1}^{\infty} b'_n e^{n\alpha_2}, \qquad (47)$$

if $\alpha_2 < 0$.

From (26) we see that the coefficients A_n, B_n, C_n, D_n for $n \geqq 2$ are determined from $\phi_n(\alpha_1)$, $\phi_n(\alpha_2)$, $\phi'_n(\alpha_1)$, $\phi'_n(\alpha_2)$, and similarly A'_n, B'_n, C'_n, D'_n are determined from $\psi_n(\alpha_1)$, $\psi_n(\alpha_2)$, $\psi'_n(\alpha_1)$, $\psi'_n(\alpha_2)$,

The values of $\phi_1(\alpha_1)$, $\phi_1(\alpha_2)$, $\phi'_1(\alpha_1)$, $\phi'_1(\alpha_2)$ will give four equations to determine the three constants A_1, B_1, C_1, and the condition that they shall be consistent gives one relation between B_0 and K. The values of $\psi'_1(\alpha_1)$, $\psi'_1(\alpha_2)$ determine the two constants A'_1, C'_1, and $\psi_1(\alpha_1)$, $\psi_1(\alpha_2)$ are not otherwise determined.

We have thus just sufficient equations to determine the coefficients in (28) with the exception of B_0, K, between which we have found one relation. If $\alpha_2 > 0$, so that the region considered lies entirely on one side of the axis $\alpha = 0$, we may take $K = 0$. If on the other hand $\alpha_2 < 0$ the condition that the stress shall vanish at infinity, which is $h\chi \rightarrow 0$ when $\alpha, \beta \rightarrow 0$, gives one more relation between the coefficients, so that in either case B_0, K are determined.

We may therefore adopt the following method:—Insert terms of the type (24) or (25) corresponding to the resultant force and couple on each boundary, and calculate the residual stresses over the boundaries. These will now form systems in statical equilibrium over each boundary, and we have shown how to determine an appropriate function of the form (28).

The problem of finding the appropriate stress-function for given tractions over the boundaries might have been approached by investigating the values of $h\chi$ and its normal gradient on the boundaries, on the lines developed by MICHELL.* The direct method which we have adopted is, however, in most cases simpler in our particular co-ordinates.

There is an exception to this rule, namely, when a boundary is free from stress. In this case the boundary conditions assume a very simple form. From (6) we have

$$\frac{\partial}{\partial \alpha}(h\chi) = \text{const} = \rho, \text{ say} \qquad (48)$$

and

$$(\cosh \alpha - \cos \beta)\frac{\partial^2}{\partial \beta^2}(h\chi) - \sin \beta \frac{\partial}{\partial \beta}(h\chi) + \cosh \alpha\,(h\chi) = \rho \sinh \alpha$$

the solution of which is readily found to give

$$h\chi = \rho \tanh \alpha + \sigma(\cosh \alpha \cos \beta - 1) + \tau \sin \beta \qquad (49)$$

on the boundary considered.

---

* 'Proc. London Mathematical Society,' vol. xxi., 1900, p. 100.

The relations (48) and (49) are the necessary and sufficient conditions that a boundary $\alpha =$ constant should be free from stress. The constants ρ, σ, τ are MICHELL's three constants of the boundary.

§ 5. A CYLINDER OR PIPE WITH ECCENTRIC BORE.

In this section we will consider the problem of a cylinder, whose cross-section is bounded by two non-concentric circles, which is subject to a uniform normal pressure over its internal surface and a different uniform normal pressure over its external surface. By FILON's theorem of generalised plane stress precisely the same analysis will give the average stresses in a plate of the same section under the same applied forces.

Let the boundaries of the cross-section be defined by $\alpha = \alpha_1$ for the internal boundary and $\alpha = \alpha_2$ for the external boundary. Then α_1, α_2 are positive and $\alpha_1 > \alpha_2$. Let the applied pressures be P_1, P_2 respectively, so that $\widehat{\alpha\alpha} = -P_1$ on $\alpha = \alpha_1$, $\widehat{\alpha\alpha} = -P_2$ on $\alpha = \alpha_2$ and $\widehat{\alpha\beta} = 0$ on both boundaries.

Let us assume

$$h\chi = B_0\alpha\,(\cosh\alpha - \cos\beta) + (A_1\cosh 2\alpha + B_1 + C_1\sinh 2\alpha)\cos\beta.$$

Calculating $\widehat{\alpha\alpha}$, $\widehat{\beta\beta}$, by means of (6) and applying the boundary conditions, we find the following values for the constants :—

$$B_0 = 2a\,M\,(P_1 - P_2)\cosh(\alpha_1 - \alpha_2)$$

$$A_1 = -aM\,(P_1 - P_2)\sinh(\alpha_1 + \alpha_2)$$

$$C_1 = aM\,(P_1 - P_2)\cosh(\alpha_1 + \alpha_2)$$

$$B_1 = aM\,\{P_1\cosh(\alpha_1 - \alpha_2)\sinh 2\alpha_2 - P_2\cosh(\alpha_1 - \alpha_2)\sinh 2\alpha_1 + (P_1 + P_2)\sinh(\alpha_1 - \alpha_2)\}$$

where, for brevity, we have written

$$M = \tfrac{1}{2}\operatorname{cosech}(\alpha_1 - \alpha_2)\,\{\sinh^2\alpha_1 + \sinh^2\alpha_2\}^{-1}$$

The most important aspect of the problem is the value of the stress $\widehat{\beta\beta}$ in the boundaries, for it is upon this that the strength of the cylinder will depend. This is most readily determined by (7), and we find without difficulty

$$\widehat{\alpha\alpha} - \widehat{\beta\beta} = 4M\,(P_1 - P_2)(\cosh\alpha - \cos\beta)\,\{\sinh(\alpha_1 + \alpha_2 - 2\alpha)\cos\beta - \sinh\alpha\cosh(\alpha_1 - \alpha_2)\}$$

so that on $\alpha = \alpha_1$

$$\widehat{\beta\beta} = -P_1 + 4\,(P_1 - P_2)\,M\,(\cosh\alpha_1 - \cos\beta)\{\sinh(\alpha_1 - \alpha_2)\cos\beta + \sinh\alpha_1\cosh(\alpha_1 - \alpha_2)\} \quad (50)$$

and on $\alpha = \alpha_2$

$$\widehat{\beta\beta} = -P_2 - 4\,(P_1 - P_2)\,M\,(\cosh\alpha_2 - \cos\beta)\{\sinh(\alpha_1 - \alpha_2)\cos\beta - \sinh\alpha_2\cosh(\alpha_1 - \alpha_1)\} \quad (51)$$

In order to investigate these results further we will consider separately the cases when the cylinder is subject to *either* internal *or* external pressures. There is no greater difficulty in the consideration of the general case, should the necessity arise, except that the formulæ are correspondingly longer.

A Cylinder under Internal Pressure.

If we put $P_2 = 0$, we have on the external surface

$$\widehat{\beta\beta} = -4P_1 M(\cosh\alpha_2 - \cos\beta)\{\sinh(\alpha_1-\alpha_2)\cos\beta - \sinh\alpha_2\cosh(\alpha_1-\alpha_2)\},$$

if d_1, d_2 denote the distances of the circles α_1, α_2 from the origin, r_1, r_2 their radii and d the distance apart of their centres, so that $d = d_2 - d_1$ we may show from (2) that

$$d_1 = a\coth\alpha_1, \qquad d_2 = a\coth\alpha_2$$

$$r_1 = a\,\mathrm{cosech}\,\alpha_1, \qquad r_2 = a\,\mathrm{cosech}\,\alpha_2$$

and

$$d_1 = (r_2^2 - r_1^2 - d^2)/2d, \qquad d_2 = (r_2^2 - r_1^2 + d^2)/2d$$

$$a^2 = \{r_2^2 - (r_1+d)^2\}\{r_2^2 - (r_1-d)^2\}/4d^2.$$

By means of these relations we can reduce the expression for $\widehat{\beta\beta}$ to the form

$$\widehat{\beta\beta} = \frac{2P_1 r_1^2\{r_2^2(r_2 - 2d\cos\beta)^2 - (r_1^2-d^2)^2\}}{(r_1^2+r_2^2)\{r_2^2-(r_1+d)^2\}\{r_2^2-(r_1-d)^2\}}.$$

From this and the obvious inequality $d < r_2 - r_1$ we easily see that—

(1) The numerically greatest stress is when $\beta = \pi$, *i.e.*, on the line of centres at the thinnest part of the cylinder. This is always a tension if P_1 is positive and is given by

$$\frac{2P_1 r_1^2(r_2^2+r_1^2+2r_2d-d^2)}{(r_1^2+r_2^2)(r_2^2-r_1^2-2r_2d+d^2)} \quad \ldots\ldots\ldots \quad (52)$$

(2) If the centre distance is greater than half the external radius there is minimum stress at the points corresponding to $\cos\beta = r_2/2d$. This is always negative when P_1 is positive and we have maximum compressions equal to

$$\frac{2P_1 r_1^2(r_1^2-d^2)^2}{(r_1^2+r_2^2)\{r_2^2-(r_1+d)^2\}\{r_2^2-(r_1-d)^2\}}. \quad \ldots\ldots\ldots \quad (53)$$

This is always numerically less than the maximum tension. There is a secondary maximum at $\beta = 0$, *i.e.*, on the line of centres at the thickest part of the cylinder, which is equal to

$$\frac{2P_1(r_2^2+r_1^2-2r_2d-d^2)}{(r_1^2+r_2^2)(r_2^2-r_1^2+2r_2d+d^2)} \quad \ldots\ldots\ldots \quad (54)$$

(3) If the centre distance is less than half the external radius, we have, in addition to the maximum tension (52), a minimum at $\beta = 0$ given by (54). There are no other maxima or minima and the stress decreases steadily from its value at the thinnest part of the cylinder to its value at the thickest part.

On the internal surface we have

$$\widehat{\beta\beta} = -\mathrm{P}_1 + 4\mathrm{P}_1\mathrm{M}\,(\cosh\alpha_1 - \cos\beta)\,\{\sinh(\alpha_1-\alpha_2)\cos\beta + \sinh\alpha_1\cosh(\alpha_1-\alpha_2)\},$$

or, expressed in terms of the radii and centre distance,

$$\widehat{\beta\beta} = -\mathrm{P}_1 + \frac{2\mathrm{P}_1 r_2^2\{(r_2^2-d^2)^2 - r_1^2(r_1+2d\cos\beta)^2\}}{(r_1^2+r_2^2)\{r_2^2-(r_1-d)^2\}\{r_2^2-(r_1+d)^2\}}. \quad . \quad . \quad . \quad . \qquad (55)$$

Hence it may be shown that

(1) If the centre distance is greater than one-half the internal radius the maximum stress in the internal surface occurs at the points corresponding to $\cos\beta = -r_1/2d$ and is

$$-\mathrm{P}_1 + \frac{2\mathrm{P}_1 r_2^2(r_2^2-d^2)^2}{(r_1^2+r_2^2)\{r_2^2-(r_1-d)^2\}\{r_2^2-(r_1+d)^2\}} \quad . \quad . \quad . \quad . \quad . \qquad (56)$$

(2) If the centre distance is less than one-half the internal radius the maximum stress is at $\beta = \pi$, *i.e.*, on the line of centres at the thinnest part of the cylinder. It is

$$-\mathrm{P}_1 + \frac{2\mathrm{P}_1 r_2^2(r_2^2+r_1^2-2r_1d-d^2)}{(r_1^2+r_2^2)(r_2^2-r_1^2-2r_1d-d^2)}. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (57)$$

(3) The minimum stress is at $\beta = 0$, the point where the line of centres meets the internal boundary at the thickest part of the cylinder. It is

$$-\mathrm{P}_1 + \frac{2\mathrm{P}_1 r_2^2(r_2^2+r_1^2+2r_1d-d^2)}{(r_1^2+r_2^2)(r_2^2-r_1^2+2r_1d-d^2)}. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (58)$$

This may be shown to be essentially positive if P is positive so that, as would be expected, the internal boundary is everywhere in a state of tension.

A Cylinder under External Pressure.

Putting $\mathrm{P}_1 = 0$ in (50) and (51) we have on the internal surface

$$\begin{aligned}\widehat{\beta\beta} &= -4\mathrm{P}_2\mathrm{M}\,(\cosh\alpha_1 - \cos\beta)\,\{\sinh(\alpha_1-\alpha_2)\cos\beta + \sinh\alpha_1\cosh(\alpha_1-\alpha_2)\}\\ &= -\frac{2\mathrm{P}_2 r_2^2\{(r_2^2-d^2)^2 - r_1^2(r_1+2d\cos\beta)^2\}}{(r_1^2+r_2^2)\{r_2^2-(r_1-d)^2\}\{r_2^2-(r_1+d)^2\}}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (59)\end{aligned}$$

and on the external surface

$$\begin{aligned}\widehat{\beta\beta} &= -\mathrm{P}_2 + 4\mathrm{P}_2\mathrm{M}\,(\cosh\alpha_2 - \cos\beta)\,\{\sinh(\alpha_1-\alpha_2)\cos\beta - \sinh\alpha_2\cosh(\alpha_1-\alpha_2)\}\\ &= -\mathrm{P}_2 - \frac{2\mathrm{P}_2 r_1^2\{r_2^2(r_2-2d\cos\beta)^2 - (r_1^2-d^2)^2\}}{(r_1^2+r_2^2)\{r_2^2-(r_1-d)^2\}\{r^2-(r_1+d)^2\}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (60)\end{aligned}$$

Hence if the centre distance is less than half the internal radius the compression in the inner surface decreases steadily from a maximum at the thinnest part of the cylinder to a minimum at the thickest part; otherwise there is a minimum at each of the points and maxima at the points corresponding to $\cos\beta = -r_1/2d$. Similarly if the centre distance is less than half the external radius the compression in the outer surface decreases steadily from a maximum at the thinnest part of the cylinder to a minimum at the thickest part; if the centre distance exceeds this value the compression is a maximum at each of these points and minima at the points corresponding to $\cos\beta = r_1/2d$.

If in these results we put $d = 0$, we have, for a concentric tube under internal pressure, tensions at the inner and outer surfaces which are respectively

$$\frac{r_2^2+r_1^2}{r_2^2-r_1^2}P_1, \qquad \frac{2r_1^2}{r_2^2-r_1^2}P_1,$$

while for a tube under external pressure the compressions at the inner and outer surfaces are respectively

$$\frac{2r_2^2}{r_2^2-r_1^2}P_2, \qquad \frac{r_2^2+r_1^2}{r_2^2-r_1^2}P_2,$$

These are the well-known formulæ for thick tubes.

§ 6. A SEMI-INFINITE PLATE WITH A CIRCULAR HOLE SUBJECT TO A UNIFORM NORMAL PRESSURE.

If in the results of the last section we put $\alpha_2 = 0$ and $P_2 = 0$, we have the solution for a semi-infinite plate containing a circular hole, which is subject to a uniform normal pressure, and bounded by a straight edge which is free from stress.

We have on the boundary of the hole

$$\widehat{\beta\beta} = -P_1+2P_1 \operatorname{cosech}^2\alpha_1(\cosh^2\alpha_1-\cos^2\beta)$$

and on the straight edge

$$\widehat{\beta\beta} = -2P_1 \operatorname{cosech}^2\alpha_1(1-\cos\beta)\cos\beta.$$

If r is the radius of the hole, d the perpendicular distance of its centre from the straight edge, and x the distance measured along the straight edge from the foot of the perpendicular,

$$d = a\coth\alpha, \qquad r = a\operatorname{cosech}\alpha_1, \qquad d^2-r^2 = a^2,$$

and

$$x = a\sin\beta/(1-\cos\beta).$$

We have therefore on the straight edge

$$\widehat{\beta\beta} = -4P_1\frac{r^2(x^2-d^2+r^2)}{(x^2+d^2-r^2)^2}. \qquad (61)$$

This has a maximum tension at the symmetrical point ($x = 0$) of magnitude

$$4P_1r^2/(d^2-r^2). \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (62)$$

At the points $x = \pm\sqrt{(d^2-r^2)}$ it vanishes, and then becomes a compression which reaches a maximum value at points at distances $\pm\sqrt{3}\,(d^2-r^2)$ on either side of the foot of the perpendicular from the centre of the hole, which is numerically equal to one-eighth of maximum tension.

The stress round the circular hole may be represented by a simple geometrical construction. If in fig. 3 the centre of the circular hole is C, Q is any point on the circle, and CA the perpendicular drawn from C to the straight edge, and if Q denote the angle QAC, we easily see that

$$\tan\phi = \sin\beta \operatorname{cosech}\alpha_1,$$

and the stress round the circular hole is

$$\widehat{\beta\beta} = P_1\,(1+2\tan^2\phi). \quad . \quad . \quad . \quad . \quad (63)$$

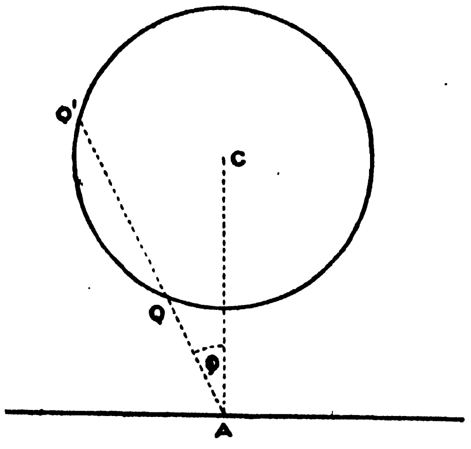


Fig. 3.

Hence the stress is the same at points Q, Q′ which lie on the same ray through A. The stress is minimum at the points nearest to and most remote from the straight edge, where it is a tension P numerically equal to the applied pressure. Thus at these points the stress is the same as it would be in the absence of the straight boundary if the plate were infinite. The maximum stresses are at the points of contact of the tangents drawn from A the circular boundary. At these points its value is

$$P_1\frac{d^2+r^2}{d^2-r^2}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (64)$$

The maximum tension in the circular boundary is equal to the maximum tension in the straight edge if $d = \sqrt{3}r$. In this case each is equal to $2P_1$. If the distance of the hole from the straight edge is greater than this value the maximum tension is at a point on the circular boundary; and if it is less, the maximum stress tension is at the symmetrical point on the straight edge. On the other hand, the point of maximum difference of principal stresses is on the straight edge or the circular boundary, according as d is greater or less than $\sqrt{2}r$. This suggests a simple method of determining whether, for a particular material, rupture occurs at the point of greatest tension or at the point of greatest stress-difference. If a circular hole is bored near the straight edge of a uniform plate, so that the distance of its centre from the edge is greater than $\sqrt{2}$ and less than $\sqrt{3}$ of the radius, and a uniform radial pressure is exerted over the hole in any convenient way and increased until rupture occurs, the crack will begin on the straight edge according to the greatest tension theory, and on the edge of the hole if the greatest stress-difference theory holds.

It will be noted that the stresses produced will become large if the hole is near to the straight edge. The formulæ are so simple that it is hardly worth tabulating their numerical values, but a single example will serve as an illustration. If the shortest distance from the hole to the straight edge is one-tenth of the radius of the hole, the maximum tension in the straight edge is 19·5 times the pressure in the hole.

§7. A SEMI-INFINITE PLATE CONTAINING AN UNSTRESSED CIRCULAR HOLE AND UNDER A UNIFORM TENSION PARALLEL TO ITS STRAIGHT EDGE.

Let the circular boundary be defined by $\alpha = \alpha_1$, so that if r is its radius and d the distance of its centre from the straight edge,

$$r = a \operatorname{cosech} \alpha_1, \qquad d = a \coth \alpha_1, \qquad d/r = \cosh \alpha_1.$$

At a distance from the hole the stress-function may be taken as $\chi = \frac{1}{2}\mathrm{T}y^2$ where T is the tension, so that, if $\alpha > 0$,

$$\begin{aligned} h\chi_0 &= \tfrac{1}{2}a\mathrm{T} \sinh^2\alpha/(\cosh \alpha - \cos \beta) \\ &= \tfrac{1}{2}a\mathrm{T} \sinh \alpha \left(1 + 2 \sum_{n=1}^{\infty} e^{-n\alpha} \cos n\beta\right) \quad . \quad . \quad . \quad . \quad . \quad . \quad (65) \end{aligned}$$

We have to add to this a stress-function which gives no stress at infinity and no stress over $\alpha = 0$, and is such that the complete stress-function gives no stress over $\alpha = \alpha_1$.

We may omit the term in K in (28), since in this case the region considered lies entirely on one side of $\alpha = 0$, and clearly the required stress-function is even in β. It may readily be seen that the condition that $\widehat{\alpha\alpha}$ and $\widehat{\alpha\beta}$ shall vanish over $\alpha = 0$ is satisfied by (28) if $\phi_n(0) = 0$ and $\phi'_n(0) = 0$ for $n \geqq 1$, and hence from (26) and (27) $\mathrm{A}_n + \mathrm{B}_n = 0$ and $(n+1)\,\mathrm{C}_n + (n+1)\,\mathrm{D}_n = 0$. We may therefore take for our complete stress-function

$$\begin{aligned} h\chi = a\mathrm{T}\Bigg[&\tfrac{1}{2} \sinh \alpha \left\{1 + 2 \sum_{n=1}^{\infty} e^{-n\alpha} \cos n\beta\right\} + \mathrm{B}_0\alpha\,(\cosh \alpha - \cos \beta) + \mathrm{A}_1\,(\cosh 2\alpha - 1) \cos \beta \\ &+ \sum_{n=2}^{\infty} \left\{\begin{array}{l} \mathrm{A}_n\,[\cosh (n+1)\,\alpha - \cosh (n-1)\,\alpha] \\ +\mathrm{E}_n\,[(n-1) \sinh (n+1)\,\alpha - (n+1) \sinh (n-1)\,\alpha] \end{array}\right\} \cos n\beta\Bigg] \quad . \quad . \quad (66) \end{aligned}$$

At infinity $\alpha = 0$, $\beta = 0$ the first series diverges, but may of course be replaced by the alternative form in (65). If the second series converges it is clear that at infinity $\chi = \chi_0$.

We must now choose the coefficients in (66) so as to satisfy (48) and (49), and there is no difficulty in finding the following values for the coefficients :—

$$A_1 = \tfrac{1}{2}e^{-2\alpha_1}\operatorname{sech} 2\alpha_1. \qquad B_0 = \operatorname{sech} 2\alpha_1 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (67)$$

$$A_n = -\frac{n^2 \sinh^2\alpha_1 - n \sinh \alpha_1 \cosh \alpha_1 + e^{-n\alpha_1} \sinh n\alpha_1}{2\left\{\sinh^2 n\alpha_1 - n^2 \sinh^2\alpha_1\right\}} \quad . \quad . \quad . \quad . \quad (68)$$

$$E_n = \frac{n \sinh^2\alpha_1}{2\left\{\sinh^2 n\alpha_1 - n^2 \sinh^2\alpha_1\right\}}.$$

Substituting in (66) we have for the complete stress-function,

$$h\chi = a\mathrm{T}\left[\begin{array}{l}\alpha \operatorname{sech} 2\alpha_1 (\cosh\alpha - \cos\beta) + \tfrac{1}{2}\sinh\alpha + \operatorname{sech} 2\alpha_1 \cosh(2\alpha_1 - \alpha)\sinh\alpha\cos\beta \\ + \displaystyle\sum_{n=2}^{\infty} \frac{\left\{\begin{array}{l} n \sinh\alpha_1 \sinh(\alpha - \alpha_1)\sinh n\alpha \\ \qquad - \sinh\alpha \sinh n(\alpha - \alpha_1)\sinh n\alpha_1 \end{array}\right\}}{\sinh^2 n\alpha_1 - n^2 \sinh^2\alpha_1} \cos n\beta \end{array}\right]. \quad (69)$$

We may now calculate the stress $\widehat{\beta\beta}$ in the boundaries by means of (6). We find, on the circular boundary $\alpha = \alpha_1$,

$$\widehat{\beta\beta}_1 = 2\mathrm{T}(\cosh\alpha_1 - \cos\beta)\left\{\sinh\alpha_1 \operatorname{sech} 2\alpha_1 + \sum_{n=2}^{\infty} \mathrm{M}_n \cos n\beta\right\} \quad . \quad . \quad . \quad (70)$$

where

$$\mathrm{M}_n = \frac{n(n-1)\sinh(n+1)\alpha_1 - n(n+1)\sinh(n-1)\alpha_1}{2\left\{\sinh^2 n\alpha_1 - n^2 \sinh^2\alpha_1\right\}}. \quad . \quad . \quad . \quad (71)$$

The stress in the straight boundary cannot be directly determined from (69), for it s found that the resulting series diverges for $\alpha = 0$. We can, however, find without difficulty from (66) that when $\alpha = 0$,

$$\widehat{\beta\beta}_0 = \mathrm{T}\left\{1 + (1 - \cos\beta)\sum_{1}^{\infty} \mathrm{P}_n \cos n\beta\right\} \quad . \quad . \quad . \quad . \quad . \quad . \quad (72)$$

where $\mathrm{P}_n = 4n\mathrm{A}_n$.

The series in (70) converges only slowly, unless α_1 is large, and for convenience in computation we may transform it by separating the more slowly converging part.

Let

$$\mathrm{M}_n = 2n(n \sinh\alpha_1 - \cosh\alpha_1)e^{-n\alpha_1} + \mathrm{N}_n, \quad . \quad . \quad . \quad . \quad . \quad . \quad (73)$$

and we readily obtain

$$2(\cosh\alpha_1 - \cos\beta)\sum_{n=1}^{\infty} n(n\sinh\alpha_1 - \cosh\alpha_1)e^{-n\alpha_1}\cos n\beta = 1 - \frac{2\sinh^2\alpha_1 \sin^2\beta}{(\cosh\alpha_1 - \cos\beta)^2},$$

Substituting in (70) we have

$$\widehat{\beta\beta}_1 = 2\mathrm{T}\left\{1 - \frac{2\sinh^2\alpha_1\sin^2\beta}{(\cosh\alpha_1 - \cos\beta)^2}\right\}$$

$$+ 2\mathrm{T}(\cosh\alpha_1 - \cos\beta)\left\{\sinh\alpha_1 \operatorname{sech} 2\alpha_1 + 2e^{-2\alpha_1}\cos\beta + \sum_{n=2}^{\infty} \mathrm{N}_n \cos n\beta\right\}. \quad (74)$$

If θ is the angle between the radius to the point α_1, β and the perpendicular to the straight edge, then

$$\sin\theta = \frac{\sinh\alpha_1 \sin\beta}{\cosh\alpha_1 - \cos\beta},$$

and if α_1 is large (74) reduces to $\widehat{\beta\beta_1} = T(1+2\cos 2\theta)$, which agrees with the known result for a hole in an infinite plate and gives compression numerically equal to T at the extremities of the diameter parallel to the tension, and tensions equal to 3T at the extremities of the perpendicular diameter.

The numerical values of the coefficients P_n, N_n are given in Tables I. and II. respectively. It will be noted from Table I. (that, as α_1 increases, P_2 tends to become

TABLE I.

α_1.	0·6.	0·8.	1·0.	1·2.	1·4.	1·6.	1·8.	2·0.	2·2.	2·4.
P_1	0·3327	0·1567	0·0719	0·0327	0·0147	0·0066	0·0030	0·0013	0·0006	0·0003
$-P_2$	3·5861	2·0401	1·2545	0·7987	0·5180	0·3400	0·2247	0·1493	0·0994	0·0664
$-P_3$	2·2393	1·0622	0·5110	0·2448	0·1160	0·0543	0·0251	0·0115	0·0053	0·0024
$-P_4$	1·3557	0·4874	0·1699	0·0570	0·0185	0·0059	0·0018	0·0006	0·0002	0·0001
$-P_5$	0·7602	0·1970	0·0474	0·0108	0·0024	0·0005	0·0001			
$-P_6$	0·3964	0·0713	0·0116	0·0018	0·0003					
$-P_7$	0·1934	0·0237	0·0026	0·0003						
$-P_8$	0·0891	0·0073	3·0005							
$-P_9$	0·0391	0·0022	0·0001							
$-P_{10}$	0·0165	0·0005								
$-P_{11}$	0·0067	0·0002								
$-P_{12}$	0·0027									
$-P_{13}$	0·0010									
$-P_{14}$	0·0004									
$-P_{15}$	0·0001									
$-P_{16}$	0·0001									

TABLE II.

α_1.	0·6.	0·8.	1·0.	1·2.	1·4.	1·6.	1·8.	2·0.	2·2.	2·4.
N_2	1·4649	0·7716	0·4139	0·2240	0·1219	0·0665	0·0364	0·0199	0·0109	0·0060
N_3	0·7457	0·2647	0·0914	0·0306	0·0100	0·0032	0·0010	0·0003	0·0001	
N_4	0·3238	0·0719	0·0148	0·0029	0·0005	0·0001				
N_5	0·1232	0·0162	0·0019	0·0002						
N_6	0·0421	0·0032	0·0002							
N_7	0·0131	0·0005								
N_8	0·0038	0·0001								
N_9	0·0010									
N_{10}	0·0003									
N_{11}	0·0001									

large compared to the other coefficients. Hence, when the hole is at a considerable distance from the straight edge, the stress in the straight edge approximates to

$$T\{1-C\cos 2\beta(1-\cos\beta)\}$$

where C is a small positive constant.

This shows that the stress in the straight edge is a minimum at the mid-point, increases to a maximum as we move outwards, then diminishes to a second minimum, and finally increases steadily to the value T at infinity, where $\beta = 0$.

In fig. 4 we have plotted the graphs of the stresses in the boundaries for a case in which the hole is fairly near to the straight edge, $\alpha_1 = 0{\cdot}8$, for which the shortest distance between the two boundaries is approximately one-third of the radius of the circle. It will be noted that the general character of the stresses is not affected by the proximity of the straight edge. It will be remembered that when the hole is at a great distance from the straight edge there are maximum stresses of 3T at the extremities of the diameter perpendicular to the straight edge, with points of maximum compression numerically equal to T lying between. For $\alpha_1 = 0{\cdot}8$ we find that the maxima occur at the same places but are increased, the increase being more marked at the point nearest to the straight edge, where the tension is $4{\cdot}366$T, while its value at the point most remote from the straight edge is $3{\cdot}266$T. The stress in the straight edge also maintains the same general character as it exhibits when the hole is at a great distance

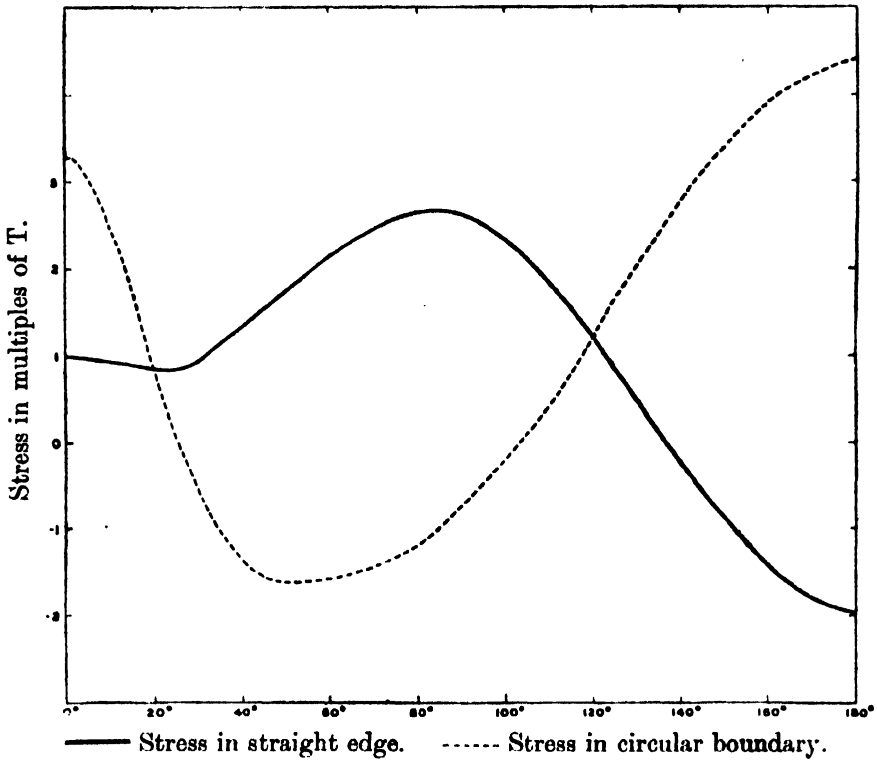


Fig. 4.

from the straight edge. Here again, however, the maxima and minima are accentuated. The minimum at the central point has decreased and has become a compression numerically equal to 1·956T.

It appears that for the range of values which we have investigated the maximum stress is on the circular boundary at the point of nearest approach to the straight edge. Its value for different values of the ratio of the distance of the centre of the hole from the straight edge to the radius of the hole, together with the stresses at the centre of the straight edge and at the most remote point of the circular boundary, is shown in Table III.

TABLE III.

α_1.	Ratio of distance of centre from edge to radius of hole.	Stress at mid-point of straight edge.	Stress at nearest point of circular boundary.	Stress at most remote point of circular boundary.
0·6	1·185	−4·080T	5·064T	3·362T
0·8	1·337	−1·956	4·366	3·266
1·0	1·543	−0·895	3·919	3·201
1·2	1·811	−0·269	3·609	3·152
1·4	2·151	+0·134	3·396	3·115
1·6	2·577	0·405	3·254	3·087
1·8	3·107	0·591	3·162	3·065
2·0	3·762	0·721	3·103	3·048
2·2	4·568	0·810	3·065	3·035
2·4	5·557	0·871	3·043	3·025
∞	∞	1·000	3·000	3·000

It will be noted that when the hole is very near to the straight edge, so that the two boundaries are separated only by a narrow connecting piece, the stress in this piece consists of a very large tension on the inside and a numerically slightly less compression on the outside. Hence, as might be expected from general considerations, the stress in this narrow connecting piece is a bending moment accompanied by a certain amount of tension.

These results may be compared with some experimental results recently obtained by Prof. COKER and Messrs. K. C. CHAKKO and Y. SATAKE* by optical means. These deal with the stresses in a strip of *finite* width under tension with a circular hole centrally placed, whereas we have considered the case of a semi-infinite plate with a circular hole near its straight edge. The problems are therefore not quite comparable; but as in each case the critical region will clearly be near the minimum section between the hole and a straight boundary, the two problems may be expected to exhibit the same general characteristics. For the strip of finite width it is found that there is maximum stress in the circular boundary at the points of nearest approach to the straight edges and minimum stress at the points of the straight edge immediately opposite the centre of the hole. Moreover as the radius of the hole is increased in proportion to the distance of its centre from the edges of the strip these maxima and minima become more pronounced. In all the cases examined experimentally the minimum stress in the straight edge remains a tension, but Prof. COKER surmises that if the radius of the hole were still further increased in proportion to the width of the strip this minimum stress would become a compression. All these results agree qualitatively with the theoretical results established in this paper for the semi-infinite plate, and allowing for the difference in the two problems they may be taken as a substantial experimental verification.

---

* 'Transactions of the Institution of Engineers and Shipbuilders in Scotland,' vol. lxiii., Part I., p. 33, 1919.

X. *The Aerodynamics of a Spinning Shell.*

By R. H. FOWLER, E. G. GALLOP, C. N. H. LOCK *and* H. W. RICHMOND, *F.R.S.*

Received January 14,—Read May 6, 1920.

CONTENTS.

Page
Introduction 296

PART I.—A GENERAL DESCRIPTION OF THE MOTION OF A SPINNING SHELL AND THE PRINCIPAL EXPERIMENTAL RESULTS.

§ 1.0. The classical theory of the plane trajectory 299
§ 1.1. The detailed specification of the complete force system 301
§ 1.2. The numerical results of the experiments 307
§ 1.3. A description in general terms of the angular motion of the axis of a shell 311

PART II.—DETAILS OF THE EXPERIMENTAL ARRANGEMENTS AND MATERIAL.

§ 2.0. General arrangements 315
§ 2.1. Measurement of the holes in the cards 319
§ 2.2. Determination of the dynamical constants 323
§ 2.3. Possible disturbing factors 324

PART III.—METHODS OF OBTAINING AND SOLVING THE EQUATIONS OF MOTION OF A SPINNING SHELL.

§ 3.0. Introductory 325
§ 3.1. The general vector equations of motion 326
§ 3.2. Equations of motion of type α 329
§ 3.3. Equations of motion of type β 331
§ 3.4. Equations of motion of type γ 333
§ 3.5. The additional force components H, J and K 334
§ 3.6. The approximate solution of equations of type α 336
§ 3.7. The solution of equations of type γ 343

PART IV.—ANALYSIS OF THE EXPERIMENTAL RESULTS.

§ 4.0. Equations of motion in polar co-ordinates 344
§ 4.1. Analysis of the experimental results 347
§ 4.2. Determination of the motion of the shell in space 357
§ 4.3. Estimate of the errors in the various solutions 362

PART V.—SUMMARY AND CONCLUSION.

§ 5.0. Summary of preceding results 367
§ 5.1. Problems for further discussion 368
§ 5.2. Effect of size and shape of shell 369

[Published December 3, 1920.

INTRODUCTION.

THIS paper contains the results, theoretical and experimental, of work undertaken, at the request of the Ordnance Committee, by the authors as Technical Officers of the Munitions Inventions Department. Permission to publish such parts as appear to be of general scientific interest has now been granted by the Ordnance Committee and the Director of Artillery. The publication of this paper has received their sanction.

The experiments in question were carried out at the firing ground of H.M.S. "Excellent," Portsmouth; the Experimental Department, H.M.S. "Excellent," also provided the 3-inch guns used and the material for the construction of the range. The authors' best thanks are due to the officers of this department, especially Lieut.-Commander R. F. P. MATON, O.B.E., R.N., without whose cordial co-operation these experiments could never have been carried out; also to the other officers of the Munitions Inventions Department who assisted in the heavy work of making and analysing the observations. The aeronautical measurements at low velocities, required for comparison, were made in the wind channels of the National Physical Laboratory, by arrangement with the Director and the Superintendent of the Aeronautical Department, to whom also we wish to express our thanks.

The subject of this paper is the motion of a spinning shell through air at velocities both greater and less than the velocity of sound. We first attempt to describe the motion of the spinning shell, considered as a rigid body, under the effects of gravity and the reaction of the air; this latter is supposed to be known in terms of the position and velocity co-ordinates of the shell, and the state of the air through which it moves. We are thus concerned throughout with the "aerodynamical" problem of the motion of the shell alone, and not with the general "hydrodynamical" problem of the motion of the complete system formed by the shell and air together. The motion of the shell thus described is then compared with the results of experiments, and the more important components of the force system imposed by the air are determined numerically as functions of certain variables such as the velocity of the centre of gravity of the shell. The actual experiments consist of observations of the initial motion of the shell (more particularly the angular motion of its axis of symmetry), over a limited range near the muzzle of the gun. The velocities experimented with range from 40 f.s.* to 2300 f.s., that is from about 0·04 to 2·1 times the velocity of sound. Using the values of the components so determined, the actual motion of the shell can be calculated with equal certainty in the more general cases which are inaccessible to direct and detailed observation.

As stated above, we make no attempt to attack the hydrodynamical problem. Such an attack is probably not yet feasible. By obtaining, however, an accurate descriptive knowledge of the force system imposed by the air, and the allied system of pressure distribution over the surface of the shell, material is provided on which a successful attack on the hydrodynamical problem may some day be based. A first contribution to a knowledge of the force system is made by the present paper. It is hoped to make a similar contribution to a knowledge of the pressure distribution in another place.*

---

* This velocity was obtained in a wind channel, using a current of air and stationary shell. The lowest velocity used in actual firing experiments was 880 f.s.

The problem proposed for discussion is of course by no means novel.† In the earlier work which is summarised by CRANZ (*cf.* (4)) the treatment of the equations of motion is often open to criticism, in view of the lack of sufficient justification for the necessary simplifying approximations. The classical theoretical results, such as MAYEVSKI's equation for the *drift*‡ (see § 4.2, equation (4.204)) have therefore hitherto justly commanded little confidence. The discussion, moreover, is of necessity based on *a priori* assumptions as to the nature of the complete force system. Unless the results of these assumptions are brought to the test of detailed experiment, the assumptions themselves must remain unjustified and unjustifiable. It should be stated here that the theory and experiments described in this paper confirm the classical theoretical results. CRANZ's own experiments (*cf.* (5)) were expressly designed to explode the fallacy that the axis of the shell, in steady motion, *precesses right round* the direction of motion of its centre of gravity. In this they are successful, but they were only carried out at low velocities, and give little in the way of quantitative results. The only real comparison of theory with experiment, which has hitherto been made, is the comparison of the observed and calculated§ values of the drift. But the observed drift is the integrated result of the disturbing forces over a considerable arc of the trajectory, and moreover, can only be disentangled with difficulty from the effect of any cross wind that may be blowing. The observed drift does not therefore serve to determine the force system with any success, though it may be used to check the values of the components otherwise determined (§ 4.21).

It may, therefore, be stated in general terms that, up to the present, there is no knowledge of the force system acting on any shell at high velocities, except when the shell is moving "nose on," *i.e.*, when its axis of symmetry and the direction of motion of its centre of gravity coincide.*

---

* "The Pressure Distribution on the Head of a Shell Moving at High Velocities," 'Roy. Soc. Proc.,' A, Vol. XCVII., p. 202.

† See for example :—

P. CHARBONNIER. (1) 'Traité de Balistique Extérieure,' ed. 2, Bk. V., Ch. IV. (2) 'Balistique Extérieure Rationnelle,' vol. II., Ch. IX.

C. CRANZ. (3) 'Lehrbuch der Ballistik; Aeussere Ballistik,' 1917, Ch. X. (4) 'Encyklopädie der Mathematischen Wissenschaften,' vol. IV., Part II., p. 185, Art. 18, "Ballistik." (5) 'Zeitschrift für Mathematik und Physik,' vol. XLIII., pp. 133, 169.

J. PRESCOTT. (6) 'Phil. Mag.,' Ser. 6, vol. XXXIV., p. 332.

Further references to previous authors will be found in (4), and the best account of CRANZ's own work in (5).

‡ The lateral departure of the projectile from the vertical plane containing the initial tangent to the path of the centre of gravity of the shell.

§ Actually, also, the important term in the calculated drift depends only on the ratio of two components of the force system, and not on their absolute values.

CRANZ (*cf.* (5)) and CHARBONNIER (*cf.* (2)) make little progress in the treatment of the general equations of motion. PRESCOTT (*cf.* (6)) makes an appreciable advance in the reductions of the general equations of motion to a tractable form, which is not too restricted in application, and gives an exact solution of his reduced equations in the simple case in which all the components of the impressed force system vary as the square of the velocity of the shell. We understand, also, that the problem of the initial motion of the shell has been recently treated by M. ESCLANGON and M. GARNIER, of the French Artillerie de Marine, with results that are closely analogous to ours, but we have not seen their work.

We therefore propose, in this paper, to give in Part III. a detailed account of the complete equations of motion of a spinning shell, moving through air, and to justify as far as possible the reduction of these equations to various useful approximate forms, some of which are classical. To do this, it is of course necessary to start from certain *a priori* assumptions as to the nature of the complete force system. These assumptions, which are far less restrictive than any that have hitherto been used, are carefully analysed when they are introduced. We then, in Part IV., submit the theoretical results so obtained to the test of the experiment described in Part II.; we are thus able to justify to some extent our *a priori* assumptions, and to obtain numerical results of some precision as to the more important components of the force system acting on the shell, in the general case. These numerical results, with a general description of the actual motion of a shell, will be found in Part I.

We have seen that the information to be obtained by comparison of the observed and calculated values of the drift is of very limited value. Two alternative methods are available, both of which are employed in this paper:—

(1) The complete force system on a model shell at rest in a uniform current of air may be determined by observations in a wind channel.†

(2) Certain components of the force system on a shell moving at high velocity may be deduced from the measurements of its oscillations just after leaving the muzzle.

The highest velocity obtainable at present by the first method is 80 f.s., but by means of the "square law" (see § 1.01) the results may be extended to velocities as great as 700 f.s. For higher velocities it is necessary to fall back on the second method which is the principal subject of this paper.

---

* In this case the force system has only one component of practical importance, namely, the resistance of the air, acting in the opposite direction to the relative motion of air and shell. This force component is here called *the drag*, in conformity with aerodynamical usage. The numerical values of the drag are known with fair accuracy for certain external shapes of shell and ordinary atmospheric conditions.

† For a full description of the construction of the wind channels at the National Physical Laboratory, and their use in measuring forces on model aircraft, see COWLEY and LEVY, "Aeronautics in Theory and Experiment."

For this purpose the shell is fired horizontally through a series of cards such as are used for measuring the jump* of the gun on firing. From the shape of the holes in the cards the actual motion of the axis of the shell can be reconstructed. Initial disturbances at the muzzle give rise to angular oscillations of the shell of sufficient amplitude for accurate measurement. These oscillations are very similar to those of the axis of a spinning top under gravity. If, as a first approximation, we regard the centre of gravity of the shell as constrained to move uniformly in a straight line over the range containing the cards, and ignore frictional damping forces in both cases, then the angular motion of the axis of the top and the axis of the shell are identical, provided that (1) the top and shell have the same axial spin and axial moment of inertia; (2) the transverse moment of inertia of the top about its point of support is equal to the transverse moment of inertia of the shell about its centre of gravity; and (3) the moment of gravity about the point of the top is equal to the moment of the force system on the shell about its centre of gravity.

In this approximate case the formal solution of the two problems is identical. As is explained in § 1.3, from the periods of the oscillations of the axis of the top or shell, we can deduce the moment of the disturbing couple and *vice versa.* In the same way the nature of the decay of the oscillations can be used to determine the damping forces.

In conclusion, we feel that a word of apology may be needed for the length of the introductory part of this paper. We do not here emphasise the applications to practical gunnery of the results obtained, but these are of some importance. We have, therefore, thought it desirable that the results should be presented in such a form as to be available to those who are concerned with the practical results, but who are not prepared to follow in detail the arguments of Parts III. and IV. At the same time it has been necessary to avoid statements which, without explanations, might convey little meaning to those who have not been technically concerned with ballistics and aerodynamics. It does not appear possible to achieve these objects except at the expense of a somewhat lengthy Introduction and Part I.

PART I.—A GENERAL DESCRIPTION OF THE MOTION OF A SPINNING SHELL AND THE PRINCIPAL EXPERIMENTAL RESULTS.

§ 1.0. *The Classical Theory of the Plane Trajectory.*

According to the classical theory, a shell is supposed to move in a resisting medium like a particle on which the only forces acting are gravity, and a resistance tangential to its path, depending only on the velocity of the particle and the state of the undisturbed medium. In such circumstances the path of the particle lies in a vertical plane and is called *the plane trajectory.*[*] This theory would be exact for a shell if the axis of the shell always pointed along the tangent to the path of its centre of gravity. The total reaction between the air and the shell would then, as required, take the form of a single force, called the drag, acting by symmetry tangentially to the path of the centre of gravity, and depending only on the velocity and shape of the shell and the state of the medium. The equations of motion resulting in this simple case are insoluble in finite terms for the actual law of resistance of the air; in practice they are capable of rapid numerical solution to any desired degree of accuracy, by a variety of methods of step-by-step integration, when the drag has been specified with corresponding accuracy.

---

* The angle between the axis of the bore before firing and the initial tangent to the path of the centre of gravity of the shell.

In order to specify the drag completely it is necessary to consider with some care what are the variables on which the drag for a given shell can depend to an appreciable extent. This question is, as yet, by no means settled, and a few of the more important considerations are summarised in § 1.01. This fact does not concern us here to a very serious extent; an incomplete specification of the variables on which the drag (or, in the general case, the complete force system) depends will only invalidate the results of observation when an attempt is made to apply them to widely different conditions of the state of the resisting medium, or of the motion of the shell. The validity is unaffected when the experimental conditions are approximately repeated. It may be assumed that, in this case of symmetry, a fairly adequate expression for the drag is given by the equation

$$\text{(1.001)} \qquad \mathrm{R} = \rho v^2 r^2 f_{\mathrm{R}}(v/a),$$

where R is the total drag, ρ the density of the air (or other medium), r the radius of the shell, v the velocity of the shell, and a the velocity of sound in the undisturbed medium; all these quantities, of course, are to be measured in a consistent set of units. In the numerical work in this paper the foot, pound, second system will be used.

Since $\rho v^2 r^2$ has the dimensions of a force, the function f_{R} is a numerical coefficient, independent of the system of units chosen, called the *drag coefficient.* Existing determinations of this coefficient as a function of v/a are very inadequate from a scientific point of view; satisfactory ones could now be made. We shall not be concerned here with the determination of this coefficient, whose value we shall only require roughly in the analysis of our experiments. We may therefore regard f_{R} as known for all values of the argument from 0 to 3, for shells of the particular external shapes which we use, moving through dry (or not too nearly saturated) air, whose temperature is not too widely different from 0° C.

---

[*] From the point of view of this paper, we regard the whole theory of the plane trajectory as "classical," though its adequate treatment was only evolved during the last years of the war.

1.01. *The Functional Form of the Drag Coefficient.*—A careful consideration of the possible forms of the function f_R, from the points of view of the kinetic theory of gases and the theory of dimensions, suggests that γ, l/r, and σ/r should be possible arguments of f_R, besides v/a. Here γ is the ratio of the specific heats of the gas, l is the mean free path, and σ the effective diameter of its molecules. We may, if desired, replace l/r by the more usual viscosity argument vr/ν, where ν is the kinematical coefficient of viscosity. Wind channel work on aerofoil and airscrew models shows that the argument vr/ν is of great importance at low velocities. Its effects, however, in the case of shell models seem almost to have disappeared by the time a velocity of 40 f.s. (or at any rate 75 f.s.) is reached. RAYLEIGH[*] obtains formulæ for the pressure on a piston moving in a pipe, which show the kind of way in which γ, as well as v/a, might enter into the expression for f_R. Variations of γ are, however, very small in practice. There is experimental evidence that some argument, other than v/a or γ, has an appreciable effect in practice, and that this argument is probably not the viscosity term in the ordinary sense. It is not possible to pursue the question further here, or to assemble in detail the evidence, which is to be found in various minutes of the Ordnance Committee.

So long as the stream lines of the flow remain unaltered by a change of velocity, the motion remains dynamically similar, the drag varies as v^2, and the coefficient f_R must be a constant. The drag is then said to obey *the square law*. Experiments with air screws, of high peripheral speed, appear to show that, up to values of v/a as great as 0·7, there is no serious departure from the square law once a certain minimum velocity is exceeded, above which the ordinary viscosity effects become unimportant; this appears, from all the evidence, to be the case also for shells, the minimum velocity being of the order of 50 f.s. As velocities of less than 100 f.s. may be ignored in ballistics, it is therefore customary to assume that the drag obeys the square law exactly for all velocities less than about $0\cdot 7a$. For all such velocities the stream lines of the flow will remain nearly unaltered and the motion will be dynamically similar.

Above this velocity ($0\cdot 7a$.) the effects of the compressibility of the air become rapidly of great importance, and the whole nature of the air-flow changes as a, the velocity of sound, is reached and exceeded. These effects are represented by the variation of f_R as a function of v/a. A good typical curve showing this variation is given by CRANZ.[†] Another example will be found in fig. 4.

We have so far ignored the fact that the shell is actually spinning about its axis of symmetry. There is no evidence to show that the drag, in the case of symmetry, is appreciably affected by the spin, and its neglect is probably justified.

A more important question is the legitimacy of assuming, as we have tacitly done in (1.001), that the drag does not depend appreciably on the acceleration of the shell. With regard to the acceleration at low velocities, it is known that the effect of the air is to increase the virtual mass of any body by an amount of the order of the mass of air displaced. This is an increase of the order of 1 part in 2000, and is entirely negligible. At higher velocities, and on the general question, direct experimental evidence is unfortunately lacking. It is, however, difficult to see, by theoretical reasoning, how the past history of the shell can have any *large* effect, and there is sufficient general experimental evidence that (1.001) is, on the average, an adequate representation of the drag in the case of symmetry to be certain that the past history is of little importance, except conceivably for a very limited range of velocities, for example, in the neighbourhood of a, the velocity of sound.

§ 1.1. *The Detailed Specification of the Complete Force System.*

The theory discussed in this paper treats the shell as a rigid body which is a solid of revolution, so that its axis of symmetry coincides with a principal axis of inertia.

---

[*] "Aerial Plane Waves of Finite Amplitude," 'Scientific Papers,' vol. V., or 'Roy. Soc. Proc.,' A, vol. LXXXIV. See in particular the last section of the paper.

[†] 'Encyklop. der Math. Wiss.,' vol. IV., Part II., p. 197.

It aims at determining the exact angular motion, as well as the motion of the centre of gravity. It confirms the classical theory of the plane trajectory (in accordance with the results of experiment), by showing that the divergences of the axis of the shell from the tangent to its path are generally small, but it aims, further, at determining the magnitude and effect of these divergences.

In this general case the force system to be specified is more elaborate than in the case of the classical theory. In accordance with aerodynamical usage, we call the angle between the axis of symmetry of the shell and the direction of motion of its centre of gravity *the yaw*, and denote it by δ. When the shell, regarded for the moment as without axial spin, has a yaw δ, and the axis of the shell OA and the direction of motion OP remain in the same relative positions, the force system can by symmetry be represented, as shown in fig. 1, by the following components, specified according to aerodynamical usage.

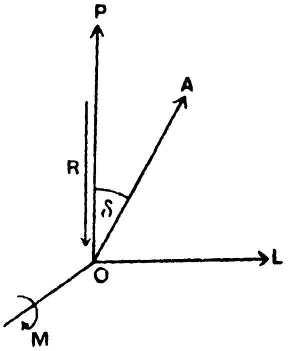


Fig. 1.

(1) *The drag*, R, acting through the centre of gravity O, in the direction of motion OP reversed.

(2) A component L, at right angles to R, called *the cross wind force*, which acts through O in the plane of yaw POA, and is positive when it tends to move O in the direction from P to A.

(3) *A moment* M about O, which acts in the plane of yaw, and is positive when it tends to increase the yaw.

By analogy with § 1.0, we assume the following forms for R, L, and M :—

$$\text{(1.101)} \qquad R = \rho v^2 r^2 f_R(v/a, \delta),$$

$$\text{(1.102)} \qquad L = \rho v^2 r^2 \sin\delta f_L(v/a, \delta),$$

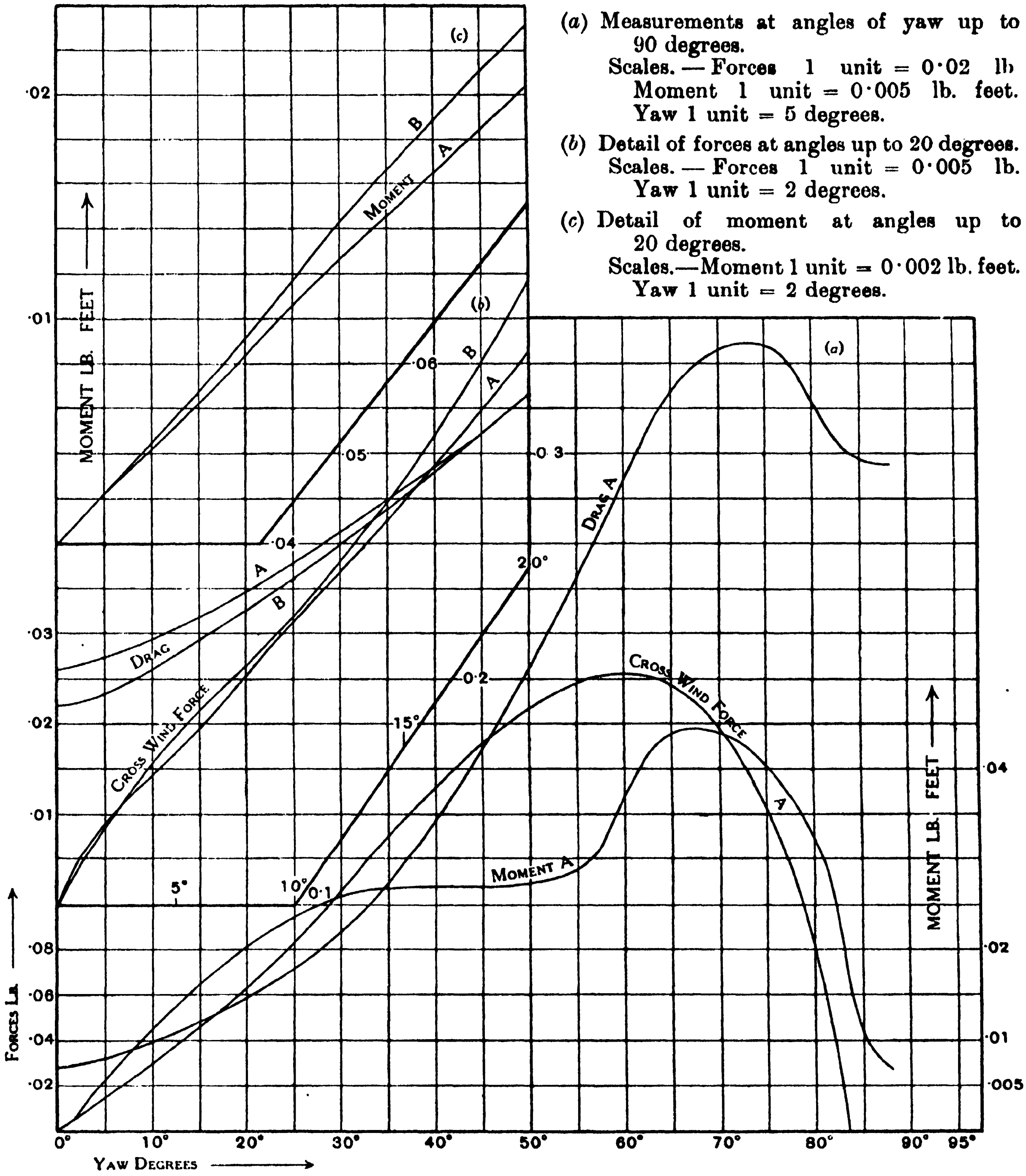
$$\text{(1.103)} \qquad M = \rho v^2 r^3 \sin\delta f_M(v/a, \delta).$$

These equations are of the most natural forms to make f_R, f_L, and f_M of no physical dimensions. The arguments of § 1.01, by which the form of equation (1.001) was justified to some extent, probably apply with equal force in this more general case. The form chosen is suggested by the aerodynamical treatment of the force system on an aeroplane. Since L and M, by symmetry, vanish with δ, the factor $\sin\delta$ is explicitly included in (1.102) and (1.103), in order that the cross wind force and moment coefficients, f_L and f_M, may have non-zero limits as $\delta \to 0$. We shall use the symbols $f_R(v/a), f_L(v/a), f_M(v/a)$ for $f_R(v/a, 0)$, $\underset{\delta\to 0}{\mathrm{Lt}}\, f_L(v/a, \delta)$, and $\underset{\delta\to 0}{\mathrm{Lt}}\, f_M(v/a, \delta)$ respectively, and shall omit the explicit mention of the argument v/a when no confusion can arise by so doing.

In view of the evidence mentioned in § 1.01, we may confidently expect that, for all values of δ, all three coefficients will be nearly independent of v/a in the region $0{\cdot}1 \leq v/a \leq 0{\cdot}7$, and shall, when required, assume their absolute independence of v/a

Fig. 2. Force components on the 3-inch shells A and B, measured in a wind channel at a wind speed of 40 f.s., plotted against angle of yaw.

Shell of form A.—Moment measured about a point 4·85 inches from base.
" " B.—Moment measured about a point 4·85 inches from base.


(*a*) Measurements at angles of yaw up to 90 degrees.
Scales.—Forces 1 unit = 0·02 lb
Moment 1 unit = 0·005 lb. feet.
Yaw 1 unit = 5 degrees.

(*b*) Detail of forces at angles up to 20 degrees.
Scales.—Forces 1 unit = 0·005 lb.
Yaw 1 unit = 2 degrees.

(*c*) Detail of moment at angles up to 20 degrees.
Scales.—Moment 1 unit = 0·002 lb. feet.
Yaw 1 unit = 2 degrees.

when $v/a \leq 0{\cdot}7$. With regard to their dependence on δ we are not here concerned experimentally with f_R. We shall assume for the purpose of analysing our experiments, where only a rough value of f_R is required, that $f_R(v/a, \delta)$ is independent of δ for small values of δ.* For the usual position of the centre of gravity of the shell, f_M at low velocities is remarkably nearly independent of δ for all values less than 10 degrees, and then diminishes as δ increases beyond this value. On the other hand, at low velocities, $f_L(v/a, \delta)$ behaves curiously for small values of δ. The wind-channel value of $f_L(v/a)$ is in consequence uncertain. Typical curves showing f_R, f_L and f_M as functions of δ at low velocities are shown in fig. 2. It is the main purpose of the experimental part of this paper to determine $f_L(v/a)$, and $f_M(v/a)$ as functions of v/a, when $v/a > 0{\cdot}7$.

1.11. *The Effect of the Angular Motion of the Axis of the Shell.*—In practice the direction of the axis of the shell relative to the direction of motion changes fairly rapidly. By analogy with the treatment of the motion of an aeroplane, we assume, tentatively, that the components of the force system R, L, and M are unaltered by the angular velocity of the axis, but that the effect of the angular motion of the axis of the shell can be represented by the insertion of an additional component, namely, a couple H, called *the yawing moment due to yawing*, which satisfies the equation

$$H = \rho v w r^4 f_H(v/a, \ldots), \tag{1.111}$$

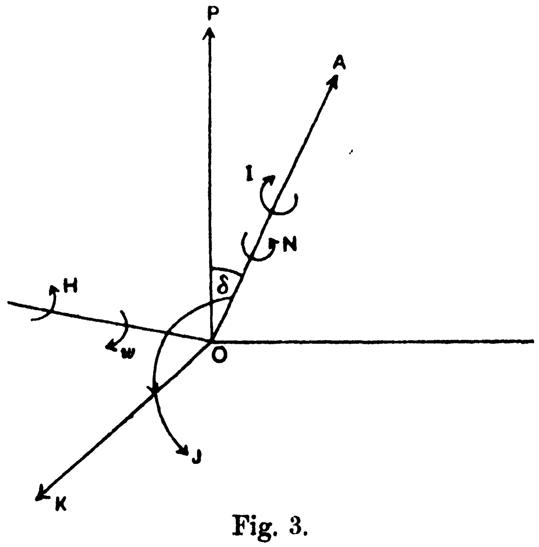


Fig. 3.

where w is the resultant angular velocity of the axis of the shell. The form of (1.111) is chosen to make f_H of no physical dimensions and is the only one suitable for the purpose. The couple H is assumed to act in such a way as directly to diminish w (see fig. 3). The yawing moment coefficient f_H may be expected to vary considerably with v/a. It may depend appreciably on other arguments such as wr/v and δ. This couple is suggested by, and is analogous to, the more important of the "rotary derivatives" in the theory of the motion of an aeroplane. It appears from considerations of symmetry that no other couple of the "rotary derivative" type need be considered. We shall arrive at rough values of f_H from our experimental results, and to some degree an *a posteriori* justification of our assumption that L and M are unaffected by the angular velocity of the axis. But the values we obtain are too rough to enable us to study the variations of f_H with any argument.

---

* By symmetry $\partial f_R/\partial\delta = 0$, when $\delta = 0$, since f_R has a minimum for $\delta = 0$. It might therefore be expected that, when δ is less than 3 degrees (say), f_R would be nearly independent of δ. This, however, is not the case in wind-channel experiments. The drag at 2 degrees and 3 degrees yaw may be 7 per cent. and 10 per cent. greater, respectively, than the drag at zero yaw. Such evidence as exists indicates that the same increase may occur also at high velocities. An experimental study of the variation of the drag with δ at high velocities would present no insuperable difficulties with modern apparatus.

1.12. *The Effect of the Axial Spin of the Shell.*—We have so far ignored the possible effect of the spin N of the shell about its axis of symmetry. We shall assume that the preceding components of the force system R, L, M and H are not appreciably affected by this spin. This is in accordance with such evidence as exists in the case of zero yaw (§ 1.01). If, moreover, the component M were seriously affected by the spin, the effect would have been detected by the present trial. No such effect was found (see § 4.13), and this fact provides some evidence of the validity of the above assumption, at least as a first approximation.

The spin N will, however, give rise to certain additional components of the complete force system. There will be a couple I which tends to destroy N, and, when the shell is yawed, a sideways force, which need not act through the centre of gravity, analogous to that producing swerve on a golf or tennis ball. This force must, by symmetry, vanish with the yaw. The swerving force must act normal to the plane of yaw, otherwise it would merely have a component which altered R or L (acting in the plane of yaw), and we have assumed that no such component exists. The complete effects of the spin N can therefore be represented by the addition to the force system of the couples I and J and the force K, acting as shown in fig. 3. To procure the correct dimensions we may assume that those components have the forms*

(1.121) $$I = \rho v N r^4 f_I,$$

(1.122) $$J = \rho v N r^4 \sin\delta f_J,$$

(1.123) $$K = \rho v N r^3 \sin\delta f_K.$$

The coefficients f_I, f_J, f_K may depend effectively on a number of variables which we can make no attempt to specify in the present state of our knowledge. These components may be expected to be very small in comparison with L and M; no certain evidence that they exist is given by our experiments.

1.13. *Relations Between the Components of the Force System.*—The various coefficients in the foregoing specifications will all depend on the external shape of the shell; results obtained for one shape cannot be applied to another. For shells of given shape, however, moving in a given manner, the forces R and L are independent of the position of O, the centre of gravity, while the moment M varies with the position of O. If M_1 and M_2 are the values of M corresponding to positions O_1 and O_2 of O, then

(1.131) $$M_1 = M_2 + O_1O_2\,(L\cos\delta + R\sin\delta),$$

where O_1O_2 is positive when O_1 is nearer the base than O_2. Using the relations (1.101) to (1.103), and assuming that the yaw is small, the equation (1.131) reduces to

(1.132) $$f_{M_1} = f_{M_2} + \frac{O_1O_2}{r}(f_L + f_R).$$

This equation is of considerable practical importance, as it enables us to deduce the values of $f_{\rm L}$ from the values of $f_{\rm M}$ for two different positions of the centre of gravity. It is found that $f_{\rm L}$ cannot conveniently be directly observed.

---

* We shall frequently write $\Gamma = I/AN$, where A is the moment of inertia about the axis of symmetry of the shell (see § 1.31).

It will be found convenient in the practical use of (1.132) to introduce the force component *normal to the axis of the shell.* If $f_{\rm N}$ is the corresponding coefficient, it is easily seen that, when the yaw is small,

$$f_{\rm N} = f_{\rm R} + f_{\rm L} \tag{1.133}$$

and that it is $f_{\rm N}$ that is directly determined by the variations of $f_{\rm M}$.

No other relations between the various coefficients are available. Previous to the present experiments, when no definite information existed as to the form of $f_{\rm L}$ and $f_{\rm M}$ as functions of v/a, special arbitrary assumptions have been made, in order to carry out calculations of the drift of a shell, or of the twisted curve described by its centre of gravity. The authors have made considerable use of the assumptions that the fractions

$$\frac{f_{\rm R}(v/a,\,\delta)}{f_{\rm R}(v/a,\,0)},\quad \frac{f_{\rm L}(v/a,\,\delta)}{f_{\rm R}(v/a,\,0)},\quad \frac{f_{\rm M}(v/a,\,\delta)}{f_{\rm R}(v/a,\,0)}$$

are independent of v/a, and have determined their values by wind-channel observations. CRANZ,[*] using essentially the same assumptions, has calculated the values of these fractions by an empirical law due to KUMMER. It must be emphasised that the use of any assumption of this type is of very dubious validity, and that, so far as experiments have yet gone, they have not confirmed any such assumptions. When the values of the coefficients $f_{\rm R}$, $f_{\rm M}$ and $f_{\rm L}$ are required for a shell of any given external shape they can and must be determined by direct experiment.

1.14. In the preceding sections, we have built up, by synthetic arguments, what appears to be the most probable complete force system. It will be seen that in so doing we have actually introduced what can be regarded as a complete system of three forces and three couples referred to three axes at right angles. Owing, however, to the complex nature of the reactions, it appears to us to be essential to construct our force system in this manner, instead of attempting to analyse a complete system of three forces and three couples, and assign each component to its proper causes. In this construction, we have been guided by considerations of symmetry, the theory of dimensions, the analogy with the theory of the aeroplane, and also, of course, by the all-important fact that the results of this construction are in agreement with experiments, so far as these have yet been carried. Of our seven components by far the most important are R, L and M; then, some way behind, H. Our experiments were designed to determine L and M, and if possible to throw some light on the size of H, and in these objects a successful start has been made. As a result, it seems reasonable to expect that the preceding specification of the complete force system will prove to be adequate; but much more work on these and other lines[†] is still required. With the numerical knowledge already obtained, which is given in § 1.2, the motion of a shell, of the shape used in these experiments, can be calculated with some approach to certainty. The general nature of the motion is described in § 1.3.

---

[*] 'Zeitschrift für Math. u. Phys.,' vol. XLIII., p. 184.

[†] For instance, the determination of the couple I that destroys the axial spin and the behaviour of $f_{\rm R}$ as a function of δ.

§ 1.2. *The Numerical Results of the Experiments.*

We now proceed to give the numerical results obtained by analysis of the observations by the methods explained in detail in Part IV.

1.21. *The Values of f_M and f_L.*—The observed values of f_M and f_L are shown plotted against v/a, in fig. 4, for the shell of external form A.* The value of f_M is

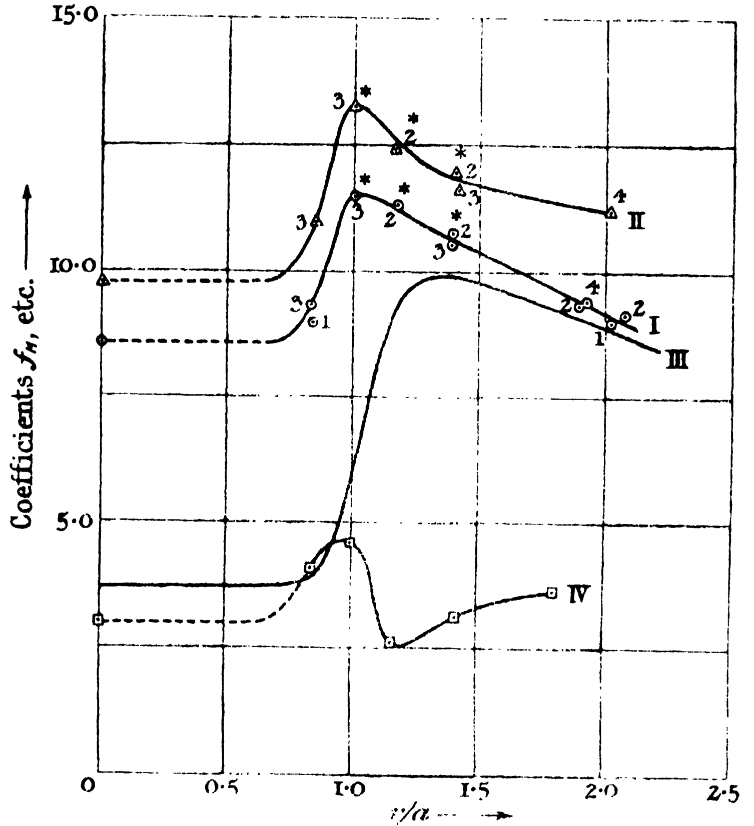


Fig. 4. Shells of form A.

Curve I.—The couple coefficient $f_M (v/a)$ for 3-inch shells, with the centre of gravity 4·73 inches from the base.

Curve II.—The same, with centre of gravity 4·20 inches from the base.

Curve III.—The drag coefficient $f_R (v/a)$ for comparison on ten times the scale.

Curve IV.—The cross-wind force coefficient $f_L (v/a)$.

The plotted points ⊙, △, ⊡ show the observed values. The numbers denote the number of observations whose mean is represented by the plotted point. The stars distinguish those groups fired from the gun rifled one turn in thirty diameters. The others were fired from a gun rifled one turn in forty.

---

* See fig. 6. Form A may be specified thus:—Length 3·84 shell diameters. Base cylindrical. Head with an ogive of 2 diameters radius. Centre of gravity 1·577 diameters from base.

given on the assumption that the centre of gravity is 4·73 inches from the base in the 3-inch shells used. The value of f_R (at zero yaw) is also given for comparison. In fig. 5, the corresponding values of f_M and f_R are given for the shells of external form B,* the centre of gravity being supposed to be 4·965 inches from the base in the 3-inch shells used. These values have been corrected as far as possible for the effect of the cards (see § 2.32), and smooth curves have been drawn through the observations. The values of f_M and f_L for shell A, and f_M for shell B, are given in the following table, Table I., for values of v/a varying by 0·1. These values have been read from the smooth curves of the figures. Besides f_L, the value of f_N, the force coefficient normal to the shell, is also given. These figures and Table I. represent the main results of the experiment. The values of f_M have a probable error of less than 2 per cent., and the values of f_L of about 10 per cent.

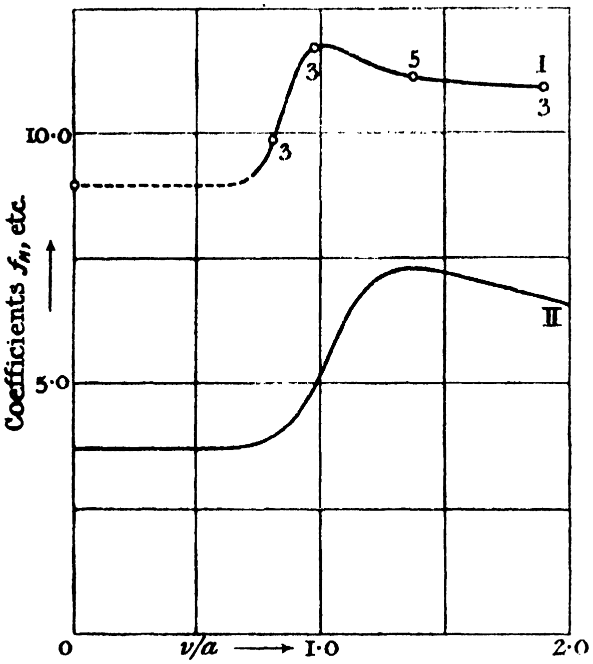


Fig. 5. Shells of external form B.

Curve I.—The moment coefficient $f_M(v/a)$ for 3-inch shells with a centre of gravity 4·965 inches from the base.

Curve II.—The drag coefficient $f_R(v/a)$ shown roughly on ten times the scale.

The differences in the various curves for f_R, f_L and f_M are very instructive. They show the complete impossibility of regarding the ratio of f_R/f_M, for example, as constant for large variations of v/a. Unlike f_R, f_M is comparatively unaffected by the velocity of sound. It increases only to about 35 per cent. above its low velocity value, and does not maintain this increase except for a narrow range of velocities near $v/a = 1$. On the other hand f_R increases to two and a-half times its low velocity value and maintains this increase.

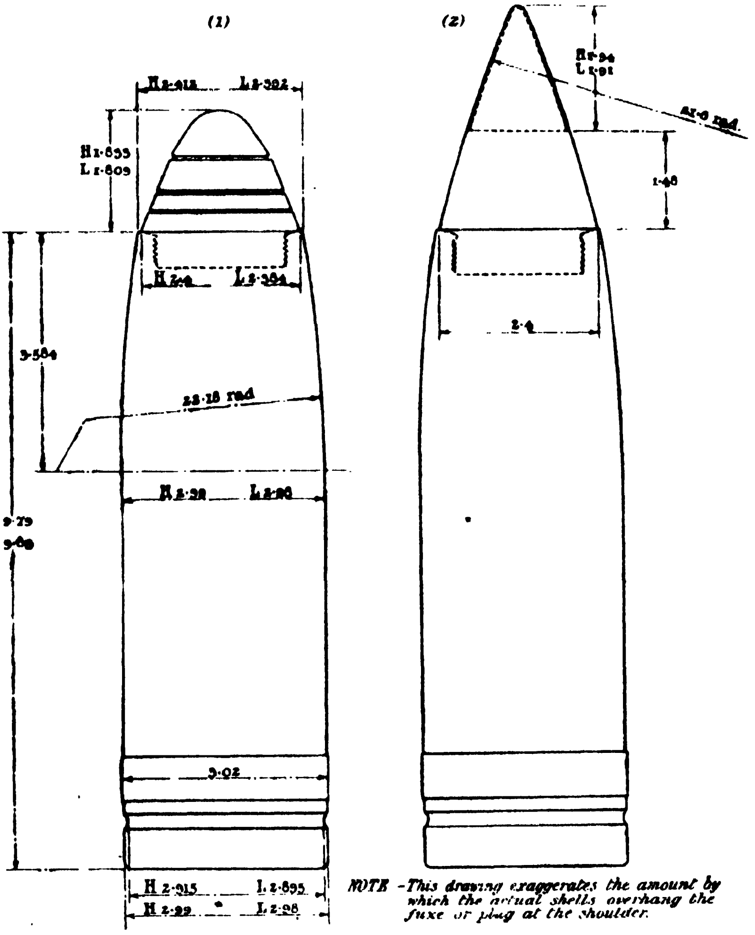
---

* See fig. 6. Form B may be specified thus:—Length 4·34 diameters. Base cylindrical. Head with an ogive of 6 diameters radius. Centre of gravity 1·655 diameters from base.

Table I.—Experimental Values of the Couple Coefficient $f_M(v/a)$, the Normal Force and Cross Wind Force Coefficients $f_N(v/a)$ and $f_L(v/a)$, for Shells of Form A, fig. 6; also Values of $f_M(v/a)$ for Shells of Form B, fig. 6.

Determined by firing trials with 3-inch shells.

v/a.	Shell of form A.			Form B.
	$f_M(v/a)$.	$f_N(v/a)$.	$f_L(v/a)$.	$f_M(v/a)$.
Wind channel.	8·57	3·34*	3·0*	8·95
0·7	8·6	—	—	9·05
0·8	9·05	4·3	3·9	9·75
0·9	10·35	—	—	11·15
1·0	11·55	5·2	4·6	11·7
1·1	11·4	—	—	11·6
1·2	11·1	3·5	2·6	11·35
1·3	10·8	—	—	11·15
1·4	10·55	4·1	3·1	11·05
1·5	10·3	—	—	11·0
1·6	10·05	4·3	3·35	11·0
1·7	9·85	—	—	10·95
1·8	9·65	4·5	3·6	10·95
1·9	9·4	—	—	10·90
2·0	9·15	—	—	—



NOTE.—This drawing exaggerates the amount by which the actual shells overhang the fuxe or plug at the shoulder.

Fig. 6. Showing the external contour of the 3-inch 16-lb. shells, Design H.E. Mark IIB, used in the trial with (1) No. 80 fuze, Mark III; (2) 6 C.R.H. plug, Design 25420.

Note.—The driving band is shown cut off at a diameter of 3·02 inches, its mean diameter after engraving.

* Uncertain.

As already mentioned in § 1.1, the low velocity value of f_L, as determined in the wind channel, is somewhat doubtful.* There appears to be a distinct minimum in this coefficient soon after the velocity of sound, followed by a steady rise. This type of curve is rather unexpected, and confirmation by a repetition of the experiment is very desirable. We may emphasise again that f_L/f_R is by no means constant, and that f_L/f_M (see fig. 15) undergoes considerable variations.

1.22. *The Force Components f_H and Others.*—We have now to exhibit the information obtainable from the damping of the oscillations of the axis. When f_L is known, this information (see § 4.12), provides numerical values for two quantities, one of which is f_H and the other $f_H+\epsilon$, where ϵ depends on the coefficients f_J and f_I and is *a priori* unlikely to be comparable with f_H. The data at our disposal are very rough and could be improved on in future experiments. The present results vary largely in some cases from round to round; the value of ϵ is much larger than its expected value and of the opposite sign. The general features of the damping (see figs. 12, 14) are however clear and qualitatively consistent. We can assert that the following rough values of f_H, given in Table II., are of the right order of magnitude and perhaps not in error by more than 50 per cent. Owing to their roughness they are given for the groups as fired. An attempt has been made to determine f_H in a wind channel at low velocities, the value 22 being obtained.

TABLE II.—Probable Values of f_H, the Coefficient of the Yawing Moment due to Yawing. Groups I., II., III. refer to shells of Form A with various positions of the centre of gravity (see § 2.2). Group IV. refers to Form B.

Group. Muzzle velocity.	f_H.	Group. Muzzle velocity.	f_H.	Group. Muzzle velocity.	f_H.
I. 22–24 1119	80	II. 24 1292	70	III. 1–4 2025	70
I. 25, 26 1326	70	II. 5–7 22, 23† 1587	75	IV. 13–15 1078	55
I. 27, 28 1563	60	II. 1–4 2024	60	IV. 16–18 1547	75
I. 1–4 2167	35	III. 17–19 1119	40	IV. 24–26 2120	80
I. 19–21 2320	30	III. 20, 21 1292	70		
II. 17–19 1119	90	III. 22, 23 1567	60		

---

* In fig. 4, and subsequently, the low velocity value of f_L is assumed to be $f_L(10°)$ in place of $\mathop{\mathrm{Lt}}_{\delta\to 0} f_L(\delta)$, which is uncertain.

† From guns of different riflings, with results in agreement.

An interesting feature of the damping is that, at a velocity of about 900 f.s., the yaw has a distinct tendency *to increase (instead of decreasing) with the time; this happens with all four types of shells.* Whether this represents a real phenomenon or is caused by the impacts on the cards (§ 4.5) is not yet clear. It is not physically impossible that f_H may be negative for this velocity. These rounds are ignored here, and further details must be postponed for Part IV.

§ 1.3. *A Description in General Terms of the Angular Motion of the Axis of a Shell.*

The numerical values of f_R, f_L and f_M described in § 1.21, with the addition of the rough values of f_H given in § 1.22 make it possible to determine numerically, by the principles of rigid dynamics, the motion of a shell projected in any manner, provided that the velocity ratio v/a, and the angle of yaw δ, do not pass outside the limits for which the determination is valid. It is necessary to obtain and solve the dynamical equations of motion in terms of the force components before proceeding to the inverse process of deducing the forces from the observed motion of the shell. Before doing so, however, it is convenient to describe in general terms the motion of the shell in various circumstances; this description is qualitative only, and is inserted for the purpose of illustration: the quantitative results are reserved for Part IV.

1.31. *The Spinning Top Analogy.*—We have already noticed in the Introduction the important analogy between the motion of the axis of a shell and the axis of a spinning top. With the reservations there made, the analogy is complete, so long as f_M can be regarded as independent of δ. The equations of motion of a stable shell, given in § 3.2, are a generalisation of the equations for the small oscillations of a top in the neighbourhood of the vertical. For the general case of stable or unstable motion where the yaw need not be small, some use can be made of the exact equations of motion of the top (§ 3.4).

In particular, the condition for *the stability* of a shell is identical with the condition for a top. The condition that the shell should be in stable equilibrium with its axis parallel to its direction of motion is that

$$(1.311) \qquad A^2N^2 > 4B\mu,$$

where A and B are, respectively, the moments of inertia of the shell about longitudinal and transverse axes through the centre of gravity, N is the spin of the shell about its (longitudinal) axis in radians per second, and $\mu \sin \delta$ is equal to M, the moment of the air forces about the centre of gravity. It is therefore convenient to define a new variable s, "the coefficient of stability," by the equation

$$(1.312) \qquad s = A^2N^2/4B\mu.$$

When s is greater than unity by a sufficiently large amount, a possible form of angular motion for both shell and top consists of a small oscillation, composed of periodic terms with two distinct periods. The values of these two periods are uniquely determined by the values of s and AN/B or Ω; conversely s and Ω, and hence μ and f_{M} are uniquely determined by the values of the periods. The main object of the jump card experiment, described in this paper, is to determine the two periods of the initial angular oscillations of a shell, fired horizontally from a gun. As Ω depends only on the spin N (known in terms of the muzzle velocity) and the moments of inertia, there is in general an independent check on the observation.* By firing the shell at a series of different muzzle velocities, values of f_{M} are determined for different values of the variable v/a, resulting in the curves of § 1.21.

1.32. The success of the experiments depends entirely on the occurrence of accidental disturbances at the muzzle, in order to produce oscillations of sufficient amplitude to be measurable. The methods of observation used were capable of giving accurate results, provided that the maximum yaw exceeded 1 degree. In the actual trial, no round was fired which developed a maximum yaw of less than 2 degrees, and it is probable that with almost any type of shell the initial disturbance would be sufficient for observations of this nature to be made. It may be noticed that, for a given initial disturbance, the amplitude of the oscillations is greater, the smaller the value of s, until, as s approaches and becomes smaller than the value unity, the amplitude of the oscillations increases very rapidly. For this reason it was at first considered preferable to deal with a shell and gun for which s was only just greater than unity, but the experiments described in this paper indicate that a value of s in the neighbourhood of $1\cdot 5$ will give the best general results.

It is to be expected *a priori*, and is confirmed by the experiment, that the initial yaw of a shell, on leaving the muzzle of a gun, is very small, and that the angular oscillations are due mainly to an initial angular velocity about a transverse axis. The shell is completely unstable under the very large pressures of the powder gases on its base, so that as soon as it is released from the barrel it is disturbed from its position of unstable equilibrium by an amount, and in a direction, which depend largely on accidental circumstances.† The pressure of the powder gases probably continue to influence the motion over a short interval after the shell has left the gun, but the whole effect on the shell must approximate to that of an impulsive couple about a transverse axis.

The angular motion of the shell, for some distance from the muzzle, approximates, therefore, to the type of motion of a spinning top known as rosette motion, in which the axis of the top passes periodically through the vertical.

---

* This check is especially important in the case of shells of type II., as the shift of the lead block on firing alters the values of the dynamical constants as determined by laboratory experiments (§ 2.2).

† [*Note added July* 31, 1920. In view of further analysis of the initial circumstances of shells in this trial, this account of the matter is probably incomplete.]

1.33. *Differences Between the Shell and Top Movements.*—We now proceed to consider the factors, so far neglected, which cause the angular motion of the shell to differ from that of the corresponding top. These may be enumerated as follows:—

(1) The effect of the cross-wind force in causing the centre of gravity to follow a curve of helical type.
(2) The effect of the force components denoted in § 1.11 and § 1.12 by H, J, and K.
(3) The effect of the diminution of forward velocity caused by the drag.
(4) The effect of gravity.

These effects will be considered in turn.

1.331. The angular oscillations of the shell give rise to a cross-wind force, which varies in magnitude and direction as the yaw varies, and this modifies the straight line motion along the direction of projection into motion of a helical type. If this helical motion could be observed with accuracy it would give valuable data for the cross-wind force coefficient f_L, but unfortunately the amplitude of the oscillations is too small to allow of this. Hence the most important effect, from the point of view of these experiments, is the reaction of the sideways motion of the centre of gravity on the angular oscillations of the shell. This helps to damp out the oscillations.

1.332. The yawing moment factor H has a similar damping effect as it is always opposed to the transverse angular velocity. While the effect of the former factor is to damp the slow period oscillation and slightly augment the quick oscillation, this latter has exactly the reverse effect. In combination, they, in general, damp out the oscillations of both periods. For the 3-inch shells, used in this trial, the yawing moment damping factor is of greater importance than the cross-wind force damping factor, and the general effect is to diminish the maximum values of the yaw, and at the same time to convert the initial rosette motion into the slower steady precessional motion.* The force component, J, due to the spin, has no appreciable effect on the angular motion, but the corresponding couple K might act as a small additional damping factor.

1.333. The head resistance or drag slowly diminishes the forward velocity, and so increases the stability factor s, by diminishing μ. The change in s diminishes the amplitude of the oscillations to a limited extent, and so assists the other damping factors.

1.334. Gravity affects the angular motion of the axis of the shell by producing curvature in the trajectory. In taking account of the gravity effect it is necessary to refer the motion to axes moving with the tangent to the trajectory (see § 3.2). The effect is quite insignificant over the range covered by the present trial, but becomes of importance at later stages of the trajectory, where it is responsible for producing the drift.*

---

* There are two possible types of steady precessional motion at constant yaw, one with a quick and the other with a slow precessional velocity.

It is convenient to illustrate this effect by considering a simple case of steady motion.

1.34. *An Illustration of the Gravity Effect.*—Let the centre of gravity of a shell be constrained to move through air at a constant speed v, in a vertical circle (fig. 7), the inclination of the path to the horizontal being θ at any instant. Thus v and $d\theta/dt$ are constant. There is a possible steady motion in which the axis OA always lies in the plane through OP perpendicular to the plane of the circle, the angle AOP (δ) being constant. The couple M tending to increase δ will also be constant, so that the contemplated motion is the same as the steady motion of a top making an angle $\frac{1}{2}\pi+\delta$ with the vertical which corresponds to the normal to the plane of the circle. The angular velocity of the axis about this normal is $-\theta'$; the value of δ as given by the ordinary formula for the steady motion of a top under these conditions† is

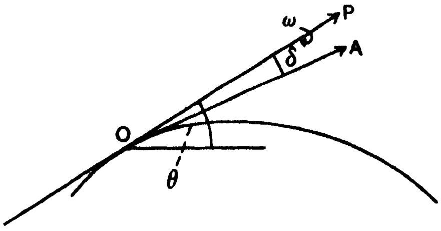


Fig. 7.

$$(1.341) \qquad -\mathrm{AN}\theta'\cos\delta+\mathrm{B}\theta'^2\sin\delta\cos\delta = \mathrm{M} = \mu\sin\delta.$$

If θ' is not too large and μ is not too small, a possible value of δ is small; we may now regard μ as independent of δ, and the equation then reduces to

$$(1.342) \qquad \delta = -\mathrm{AN}\theta'/\mu = -4s\theta'/\Omega,$$

the term neglected being of order δ^3. When a shell is moving freely the angular velocity θ' increases, and the linear velocity diminishes up to a point beyond the vertex of the trajectory. If the initial motion is identical with the above steady motion, this will cause the couple M to diminish, so that the axis of the shell will lag behind its position in the steady motion. This lag gives rise to a component angular velocity of the axis tending to increase the yaw δ, until a state of relative equilibrium is reached, in which the yaw is slightly less than its equilibrium value, and the axis lags slightly behind (*i.e.*, above) the tangent OP. When the velocity is high and the spin N not too large, M is large and the true position of the axis lies very near the equilibrium position. It will be shown in fact, in Part IV., that the assumption that OA is level with OP, and that δ is given by (1.342), lead to a determination of the drift which is sufficiently accurate for all trajectories of elevation less than 30 degrees. The drift is produced by the cross-wind force resulting from the above value of the yaw.

---

* For a shell whose spin and direction of motion are related like a right-handed screw the drift is to the right of the plane of fire.

† See, *e.g.*, ROUTH, 'Rigid Dynamics,' vol. II., Art. 207.

In the neighbourhood of the vertex of a trajectory of high elevation, both the velocity and the couple M become very small, so that δ becomes large. A calculation has been made, by a step-by-step process, of the angular motion and drift of a shell fired at an elevation of 70 degrees. The yaw, soon after the vertex, reaches the value 60 degrees, while the axis lags behind the tangent to the path by more than 45 degrees.

1.35. The effect of gravity as described in the last section completes the list of factors which have an appreciable effect on the motion, and it remains to consider the way in which they combine. It will be shown in § 3.2 that the motion of the axis of a stable shell is determined, to a good approximation, when the yaw is not too large, by a linear differential equation of the second order. The effect of gravity is to produce the type of motion described in § 1.34, given the proper initial conditions in which the yaw and its rate of increase are both very small. The complete motion under arbitrary initial conditions may be obtained by superposing the appropriate type of initial oscillatory motion, which is unaffected by gravity. The superposed oscillations will ultimately be damped out, leaving the motion of the last section only. The motion of the centre of gravity will be appreciably affected by alteration of the initial conditions only in so far as they produce a certain small sideways displacement and velocity (§ 4.2), and increase the drag to an extent which is not yet known.

More detailed results are reserved for Part IV., following the discussion of the mathematical theory. Actual examples of the observed motion of the shell's axis can be studied in fig. 14.

PART II.—DETAILS OF THE EXPERIMENTAL ARRANGEMENTS AND MATERIAL.

§ 2.0. *General Arrangements.*

We propose, in this part, to explain the details of the experiments in so far as is necessary to enable the reader to understand the method used, and to form an estimate of the accuracy obtained, or capable of being obtained, in this manner.

The experiments were carried out as the weather served in January and February, 1919, four different types of 3-inch shells being fired, at various velocities, from each of two differently rifled guns. The constants of the shells used are given in Table III.,

TABLE III.—Mean Values of the Dynamical Constants of the Shells used, Determined before Firing.

Types I., II. and III., Form A. } (See fig. 6.)
Type IV., Form B. }

Type of shell.	Length, inches.	Weight, lb.	Distance of centre of gravity from base, inches.	Axial moment of inertia, A, lb. (in.)^2.	Transverse moment of inertia, B, lb. (in.)^2.	B/A.
I. (Normal) . .	11·53	14·09	4·727	18·37	143·9	7·83(5)
II. (Centre of gravity forward)	11·53	16·31	5·124	19·20	165·0	8·59
III. (Centre of gravity back)	11·53	16·48	4·203	18·93	129·5	6·84
IV. (Shells with pointed nose)	13·15	14·62	4·965	18·71	166·2	8·89

and details of the groups fired are given in Table IV. The distance available between the firing point and the sea at Portsmouth is rather less than 600 feet. The motion of the shell was recorded over this range, within which the effects of gravity are fairly small and the path of the shell not widely different from a straight line. To achieve this the shell was fired through a series of millboard pistol targets, 2 feet square, about $\frac{1}{20}$ inch thick, which were fastened approximately at right angles to the path of the shell, at suitable distances from the muzzle.* The plane of the card was carefully adjusted, and it is probable that in no case did the angle between the path of the shell and the plane of the card differ from a right angle by as much as two degrees. As errors up to four degrees do not affect the shape and position of the hole in the card, which determine the position of the axis and the centre of gravity of the shell at the moment of impact, it may be assumed that in every case

---

* For the gun whose rifling made one complete turn in a length of 40 diameters of the bore (rifled 1 in 40) ten cards were used, placed approximately at 60-foot intervals, the first card being 50 feet from the muzzle. For the gun rifled 1 in 30 twelve cards were used, the first seven being at 30-foot intervals and the later cards at 60-foot intervals as before. The distance of the cards from the muzzle of the gun was determined with a probable error of 1 inch.

TABLE IV.—Showing Groups of Rounds Fired.

The types of shell are numbered I.–IV., and the shells of each type are numbered 1, 2, 3, ... in the order of firing.*

GUN rifled one turn in 40 diameters of the bore.

Group.	Mean muzzle velocity for group, f.s.	Remarks.	Group.	Mean muzzle velocity for group, f.s.	Remarks.
Type I.; shells of form A; centre of gravity normal.					
I. 11–14	922	Stable	I. 15, 16	2130	Stable †
I. 8–10	1072	Unstable	I. 1–4	2167	Stable
I. 17, 18	1312	Unstable	I. 19	2272	Stable
I. 5–7	1565	Just stable	I. 20, 21	2346	Stable
Type II.; shells of form A; centre of gravity forward.					
II. 8–10	934	Stable	II. 5–7	1585	Stable
II. 11–13	1107	Unstable	II. 1–4	2024	Stable
II. 14–16	1334	Unstable			
Type III.; shells of form A; centre of gravity back.					
III. 8–10	931	Stable	III. 5–7	1583	Just stable
III. 11–13	1077	Unstable	III. 1–4	2025	Stable
III. 14–16	1312	Unstable			
Type IV.; shells of form B; centre of gravity normal.					
IV. 10–12	884	Very unstable	IV. 1–6	2130	Very unstable
IV. 7–9	1553	Very unstable			

---

* Only the stable groups are analysed in this report. For a specimen yaw curve in an unstable case, see fig. 12.

† Fired with cards on the far screens only, to determine by comparison the effect of the impacts on the cards.

TABLE IV. (continued).

GUN rifled one turn in 30 diameters of the bore.

All groups were stable.

Group.	Muzzle velocity, f.s.	Group.	Muzzle velocity, f.s.	Group.	Muzzle velocity, f.s.
Type I.					
I. 22–24	1119	I. 25, 26	1326	I. 27, 28	1563
Type II.					
II. 17–19	1119	II. 24	1292	II. 22, 23	1589
Type III.					
III. 17–19	1119	III. 20, 21	1292	III. 22, 23	1567
Type IV.					
IV. 21–23	900	IV. 16–18	1547	IV. 24–26	2121
IV. 13–15	1078	IV. 19, 20*	1547		

the centre of gravity of the shell was moving normally to the card.† Thus the angle actually recorded by the shape of the hole in the card is the true yaw of the

---

* Fired with cards on the far screens only, to determine by comparison the effect of the impacts on the cards.

† The angular motion of the axis of the shell is comparatively so slow that it can be ignored during the interval in which the shell is passing through a card. For instance, with the shells used in this trial the change in ϕ, the orientation of the yaw, is never as much as 3½ degrees during the complete passage through the card, and the change in δ never as much as 8 minutes. These quantities are of the same order as the errors of observation and may be ignored. Thus the shell can correctly be regarded as equivalent for cutting purposes to its circumscribing cylinder (of indefinite length) whose generators are parallel to the direction of motion of the centre of gravity.

If the direction of motion is normal to the plane of the card at the moment of impact, a certain hole will be cut in the card, whose shape will be precisely that of the normal cross-section of this circumscribing cylinder. But if the card is tilted through a small angle τ about any axis in its own plane, the hole made by the shell will be the same as the cross-section of the supposed cylinder by the plane of the card shell, that is, the angle between the axis of the shell and the direction of motion of its centre of gravity.

On each card there was marked, by methods which need not be particularised, (*a*) the vertical, (*b*) a reference point from which the point of aim for each round could be deduced. The probable error in the marking of the vertical was negligible compared to the other errors of observation. The probable error* in each co-ordinate of the point of aim was about 0·2 inches.

Times of flight from the muzzle to each card were not directly observed, but the mean velocity of the shell over a suitable interval of the range was observed for each round with two standard Boulangé chronographs. These were sometimes used as a pair—in these cases their readings were in good agreement—and sometimes separately, at opposite ends of the range, to determine the loss of velocity, and so an approximate value for the average coefficient of the drag. From the data so obtained, the muzzle velocity and the times of flight from the muzzle to each screen were calculated by the usual ballistic methods to a nominal accuracy of 1 f.s. and 10^{-4} second, respectively. It is improbable that any of these quantities are appreciably in error to the order of accuracy required by the rest of the experiment. A check on the calculated muzzle velocity is provided by the observations, for a discussion of which the reader should refer to § 4.1.

§ 2.1. *Measurement of the Holes in the Cards.*

It is now necessary to deduce, from the position and shape of a hole in any card, the position of the axis and centre of gravity of the shell at the moment of passing the card. This can usually be done with considerable accuracy. It has been found that at all velocities less than 1600 f.s., and often at higher velocities, the hole has the form shown diagrammatically in fig. 8, and by photographs of actual examples in fig. 8A.

Inside the outer circumference ABA′B′ of the hole, a considerable amount of bruised and partly torn card QQQ is left, which is still attached to the untouched part. It is found that, when the edges of this part are flattened out, they always define with some accuracy a circle of diameter 2·40 inches. A stiff paper circle of this diameter can be fitted to the hole with such certainty that its centre is seldom in doubt by more than 0·01 or at most 0·02 inches.

---

in its tilted position. The dimensions of such a hole will only differ from those of normal impact by terms of order $d(1-\cos\tau)$, where d is any dimension of the hole. Such second-order terms are completely negligible if $\tau < 4$ degrees. Thus in all cases the shell may be regarded as cutting the hole in the card as if the direction of motion of its centre of gravity is normal to the plane of the card at the moment of impact.

* Throughout this paper "probable error" is used with its technical meaning, see *e.g.*, BRUNT, 'The Combination of Observations,' p. 30.

The external form of the shells used in the trial is shown in fig. 6. It will be observed that at the junction of the body of the shell and the fuze or plug there is a distinct cutting edge of plan diameter 2·402 inches. It is clear, therefore, that when the impact takes place, a circle of cardboard, 2·40 inches in diameter, is punched out and cleanly removed by this edge; the greater part of the circumference of this inner circle is usually removed by the subsequent passage of the body of the shell, which cuts the complete hole, but enough remains, in a bruised state, for yaws that are not too large, to define the position of the centre of this section of the shell at the moment of impact on the card.

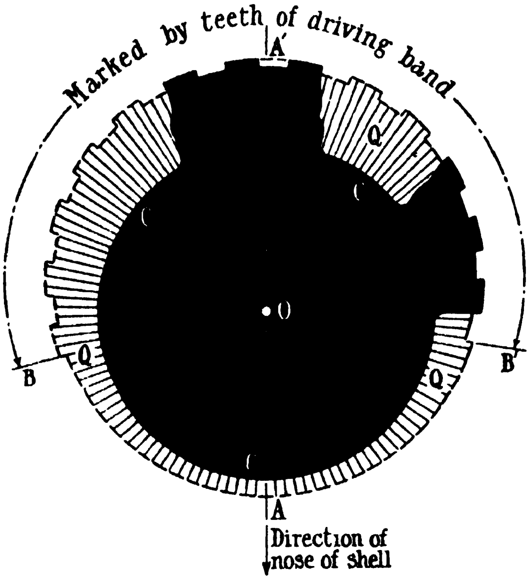


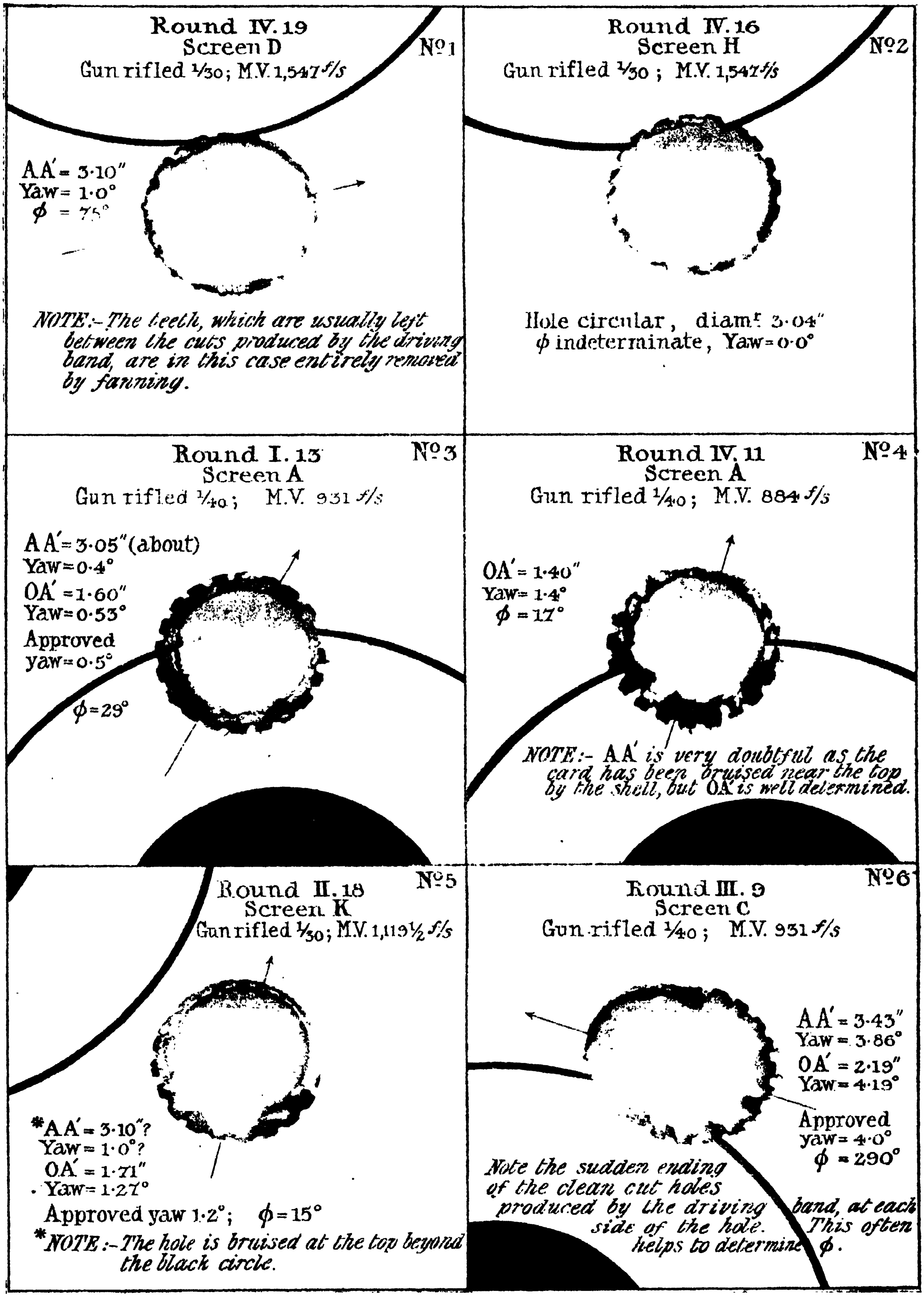
Fig. 8. Diagrammatic sketch of a typical hole, for a yaw between 1 degree and 4 degrees, when the velocity is low or medium.

CCC.	Inner circle—radius 2·40 inches, centre O.
ABA′B′.	Outer circumference of hole.
QQQ.	Bruised part of card.
AA′.	Axis of symmetry or greatest diameter of hole.
BA′B′.	Circumference cut by teeth of driving band.
BAB′.	Ditto cut by nose or shoulder of shell.

The lengths AA′ (3·16 inches in figure) and OA′ (1·80 inches) each serve to determine the size of the yaw.

The values of the yaw corresponding to the above values are 1·6 degrees and 1·8 degrees respectively, mean 1·7 degrees.

It follows, therefore, that there are two distinct methods by which the value of the yaw δ can be determined. In the first place, there is a unique relation between the greatest diameter of the hole (AA′, fig. 8) and the value of δ; secondly, there is a unique relation between OA′ and δ. These relations can be tabulated numerically when the plan dimensions of the shell are known, and the value of δ corresponding to any measured length AA′ or OA′ read off.



Round IV. 19
Screen D
Gun rifled 1/30; M.V. 1,547 f/s
No. 1

NOTE:—The teeth, which are usually left between the cuts produced by the driving band, are in this case entirely removed by fanning.

Round IV. 16
Screen H
Gun rifled 1/30; M.V. 1,547 f/s
No. 2

Hole circular, diam.r 3·04"
φ indeterminate, Yaw = 0·0°

Round I. 13
Screen A
Gun rifled 1/40; M.V. 931 f/s
No. 3

Round IV. 11
Screen A
Gun rifled 1/40; M.V. 884 f/s
No. 4

NOTE:—AA' is very doubtful as the card has been bruised near the top by the shell, but OA' is well determined.

Round II. 18
Screen K
Gun rifled 1/30; M.V. 1,119½ f/s
No. 5

**NOTE:—The hole is bruised at the top beyond the black circle.*

Round III. 9
Screen C
Gun rifled 1/40; M.V. 931 f/s
No. 6

Note the sudden ending of the clean cut holes produced by the driving band, at each side of the hole. This often helps to determine φ.

Fig. 8A.

To determine the value of ϕ,* it is necessary to measure the angle between AA′ and the vertical recorded on the card. The direction of AA′ must be determined by eye from the considerations that it is (1) the greatest diameter, (2) the axis of symmetry of the hole, and (3) that it must pass through O, which is located as the centre of a paper circle fitted into the inner hole.

By proceeding in this manner it was found that the values of δ could be nearly always determined with confidence by at least one method and often by both. When both methods were available the agreement in the resulting values of δ was in general good; the average difference between them in *all* cases for shells of type I. (99 in number), in which both measurements were available and both appeared to be *a priori* reliable, was 0·20 degrees. These cases were simply taken as a sample. The general features of the agreement were the same for all types. We may therefore fairly assert that the probable error of any determination of δ is something less than 0·2 degrees. The use of the measurement OA′ is of special importance for small values of δ, and in fact alone makes their accurate determination possible.

The probable error in the determination of ϕ is not quite so easy to estimate, as there is no alternative method of determining ϕ. The method is clearly theoretically sound, and the errors can only arise from faulty estimations of the symmetry of the hole. By making a number of independent determinations for the same hole, with proper precautions against a biassed judgment, and comparing their consistency, it appeared that the probable error of any determination of ϕ was less than 2½ degrees, unless the yaw was small (less than 0·8 degrees, say). As the yaw approaches zero, the errors in the determination of ϕ increase rapidly until, when the yaw is less than 0·2 degrees, ϕ cannot be determined at all.

Proceeding in this manner the values of δ and ϕ were tabulated for each round for the values of the time corresponding to the position of each card. If the above estimates are correct it is doubtful if the accuracy obtained could be much improved on without a radical change in the method of recording the position of the shell.

2.11. When the yaw has been determined, and the position of the centre of gravity on the axis of the shell is known, its position along AA′ can be calculated from the dimensions of the shell. The position of AA′ on the card is well determined, and so the position of the centre of gravity can be located with respect to the reference point, and so with respect to the point of aim. This part of the determination is considerably more accurate than the location of the reference point on the card. The path of the centre of gravity for a small number of rounds was measured up in this manner; the results of the discussion (§ 4.2) are mainly null, in agreement with theory. The measurements were therefore not completed for every round and are not given here.

---

* The angle ϕ denotes the angle between the plane of yaw OAP and the vertical plane through OP. See fig. 10, p. 332.

§ 2.2. *Determination of the Dynamical Constants.*

All the shells used in these experiments were weighed before firing, and their overall lengths were measured. The variations from shell to shell were small, and the mean values given in the tables may be assumed to be correct for all purposes. No appreciable change in these quantities is likely to occur on firing.

The moments of inertia were determined, before firing, for a selection of about 25 per cent. of the shells of each type. The probable error of any determination was about 1 part in 2000. The mean values for the different types of shell are given in the table. The extreme variation of any transverse moment of inertia from the mean was 1·8 per cent., and of the axial moment of inertia was 0·8 per cent. The errors in assuming that the mean value of the sample is the correct value for each round may therefore be appreciable at times, but should not seriously affect the final mean results. The general accuracy of the experiments was, contrary to expectation, sufficient to warrant the refinement of determining and using the individual values for each shell.

The centres of gravity were also determined, before firing, for the same selection of shells, and the mean value of the distance of the centre of gravity from the base is given, in the same table, for each type. The determination was made with a probable error of 0·003 inches. The values were fairly constant for the shells of any one type, the extreme variation from the mean being 0·022 inches.

It is by no means certain *a priori* that the values of A and B and the position of the centre of gravity may not be changed appreciably in some of the shells by the stresses set up when the gun is fired. No change is at all likely in the empty shells of types I. and IV., or in the bodies of the other shells; they may be confidently relied upon not to be stressed beyond their elastic limit; but the lead and wood filling in the shells of types II. and III. is decidedly suspect. To test this point, two shells of each of the types II. and III., after the determination of their dynamical constants, were fired* over water for recovery, and their constants were then re-determined. In the case of the shells of type III., with a filling of lead at the back and wood in front, there was no appreciable change. In the case of the shells of type II. with lead in front and wood behind, the wood block, as might have been expected, was crushed, and the lead had moved back about three inches in the case of the high velocity and one inch in the case of the low. The axial moments of inertia, A, were unaltered, but the transverse moments of inertia B and the positions of the centre of gravity were of course seriously affected. It was found, however, that the observed changes in both could be satisfactorily accounted for by the observed movement of the lead block, of weight 1·9 lb. When the centre of gravity of the shell of type II. is 4·727 inches from the base, so that it coincides with the centre of gravity of a shell of type I., the value of B is 145·7 lb. (inch)². Neglecting the effect of the wood, suppose that the lead plug is x inches further forward. In such a case

---

* One of each type at a muzzle velocity of 1950 f.s. and one at 1530 f.s.

$$B = 145{\cdot}7 + 1{\cdot}9x^2. \tag{2.21}$$

If, moreover, l is the distance of the centre of gravity from the base in inches, then

$$l = 4{\cdot}727 + 0{\cdot}117x. \tag{2.22}$$

The altered position of the centre of gravity can therefore be recovered by calculation, if the altered value of B can be deduced from the observations. This is, in fact, the case (see § 4.1), so that even for shells of type II. the dynamical constants of the shells after firing are satisfactorily certain.

§ 2.3. *Possible Disturbing Factors.*

There are two further possible causes of error which we have not yet mentioned. These are (1) the wind, and (2) the impulsive action between the shell and the card.

2.31. *The Effect of Wind.*—Since we are studying the motion of the shell under the force system impressed by the air, we are concerned solely with the motion of the projectile *relative to the air*, but we can only observe, by means of jump cards, the motion of the projectile *relative to the ground.*

If the strength and the direction of the wind are known, it is an easy matter to convert the observed values of the size and orientation of the yaw, and the observed motion of the centre of gravity, into the corresponding quantities for the motion relative to the air. It is, however, very difficult to determine what is the strength of the wind, at the moment of firing, only a few feet above the ground. It is, therefore, necessary to carry out jump card trials in calm weather. During the experiments the wind exceeded 10 f.s., only at the moments of firing three rounds, and was usually only 5 or 6 f.s. at 20 feet above the ground. Its strength near the ground will have been still less, and its effects may therefore be neglected.

2.32. *The Impulsive Action between the Shell and the Card.*—When the experiments were started it was not expected that the effect of the cards would be decidedly bigger than the probable random errors of the results. This, however, appears to be the case. A limited amount of evidence, for determining the necessary correction, is supplied by the few comparative rounds fired without cards on the nearer screens. Such comparative rounds would have been included in all, or at least the majority, of the groups, if their importance had been realised earlier. The evidence supplied by the comparative rounds was carefully analysed, and was supplemented, after the conclusion of the trial, by determination of the *magnitude* of the impulse between the cards and the shells by observation of the extra loss of velocity so caused. The magnitude of a single impulse, at not too great a value of the yaw, probably has the values

14·3 foot-poundals at 2470 f.s.,

8·9 foot-poundals at 1140 f.s.;

the values at other velocities may be roughly obtained by linear interpolation.

The effect on the observed motion of the axis due to an impulsive couple was calculated, and it was found that rough values could be assigned for the magnitude of the impulsive couple acting at any card. On calculating the total effect on the observed value of s it was found that the probable correction required varied from 2½ to 4½ per cent. in the various groups. This correction was applied before constructing Table I. and figs. 4 and 5. The figures of Table II. have not been corrected for this effect as their accuracy is not great enough to make it worth while to do so.

PART III.—METHODS OF OBTAINING AND SOLVING THE EQUATIONS OF MOTION OF A SPINNING SHELL.

§ 3.0. *Introductory.*

On the assumptions discussed in Part I. the equations of motion of a spinning shell can be written down at once by the rules of rigid dynamics. Three different types of these equations will be found of use in practice, all of which may be obtained most simply as special cases of the vector equations of motion of the shell, referred to axes rotating in the most general manner. The use of the vector notation, in the initial stages of the discussion, has the further advantage of showing most clearly the meaning of the various terms, and of presenting the results in a symmetrical form.

In order to simplify the general equations, the only components of the force system impressed by the air, retained in the initial discussion, are R, L, M, and the spin-retarding couple I $(= \mathrm{AN}\Gamma)$. The remaining components are of less importance and will be inserted later on in § 3.5.

After obtaining the general equations the three special types are deduced. They may be described as follows:—

Type α.—Equations in terms of direction cosines, referred to axes moving with the tangent to the corresponding plane trajectory.

Type β.—Equations in terms of direction cosines or spherical polar co-ordinates, referred to axes moving with the tangent to the actual twisted trajectory.

Type γ.—Equations similar to the equations of energy and angular momentum of a top (spherical polar co-ordinates), referred to the axes used for type β.

In each case the equations obtained are simplified by certain approximations, and the results are suitable for use only under certain conditions. Equations of type α are valid when the shell is sufficiently stable and the yaw is small; type β when the shell has settled down to a non-periodic motion in which the yaw may be large, the initial oscillations being damped out; and type γ when the motion of the centre of gravity is nearly rectilinear.

Equations of these types cannot be solved exactly, and the method of approximation used to obtain a solution is different in each case. The equations of type α are used for the analysis of the jump card experiments, for all sufficiently stable rounds, and could be used to compute the entire motion in any trajectory whose initial elevation is less than 45 degrees. Equations of type β have been used to compute the latter part of a twisted trajectory at an elevation of 70 degrees. Equations of type γ have a limited application in analysing the jump card records for rounds which are nearly or quite unstable.

3.01. *Note on the Vector Notation.*—All letters which represent vector quantities will be in clarendon type, to distinguish them from scalar quantities in the ordinary type. The three components of any vector **A**, referred to right-handed rectangular axes 1, 2, 3, are written A_1, A_2, A_3.

If **A** and **B** are two vectors, their vector product is denoted by $[\mathbf{A}\,.\,\mathbf{B}]$. This represents the vector whose components are

$$(A_2B_3-A_3B_2), \qquad (A_3B_1-A_1B_3), \qquad (A_1B_2-A_2B_1).$$

It is perpendicular to the plane containing the two vectors in the direction of the axis of the right-handed screw, which turns from **A** to **B**, its modulus being equal to the product of the moduli of **A** and **B** into the sine of the angle between them. The scalar product of the two vectors is written $(\mathbf{A}\,.\,\mathbf{B})$, and is equal to the scalar quantity

$$A_1B_1+A_2B_2+A_3B_3;$$

it is also equal to the product of the moduli of **A** and **B** int [illegible] nds, and was usually between them, being positive when this angle is acute. For [illegible] ground [illegible] $(\mathbf{A}\,.\,\mathbf{A})$ by $(\mathbf{A})^2$, which is equal to the square of the modulus of **A**.

angle
lenote

Constant use is made of the following identities:—

$$[\mathbf{A}\,.\,\mathbf{A}] = 0, \; ([\mathbf{A}\,.\,\mathbf{B}]\,.\,\mathbf{A}) = 0. \tag{3.011}$$

$$[[\mathbf{A}\,.\,\mathbf{B}]\,.\,\mathbf{C}] = (\mathbf{A}\,.\,\mathbf{C})\,\mathbf{B}-(\mathbf{B}\,.\,\mathbf{C})\,\mathbf{A}. \tag{3.012}$$

$$([\mathbf{A}\,.\,\mathbf{B}]\,.\,[\mathbf{B}\,.\,\mathbf{C}]) = (\mathbf{A}\,.\,\mathbf{B})\,(\mathbf{B}\,.\,\mathbf{C})-(\mathbf{B})^2\,(\mathbf{A}\,.\,\mathbf{C}). \tag{3.013}$$

§ 3.1. *The General Vector Equations of Motion.*

We take a system (1, 2, 3) of right-handed axes of reference, see fig. 9, whose origin is O, the centre of gravity of the shell, and whose angular velocity at any instant is represented by the vector $\boldsymbol{\Theta}$, with components Θ_1, Θ_2, Θ_3. The direction of the axis of the shell OA is represented by the unit vector* $\boldsymbol{\Lambda}$, and the direction of motion of the centre of gravity by the unit* vector $\mathbf{X}$.

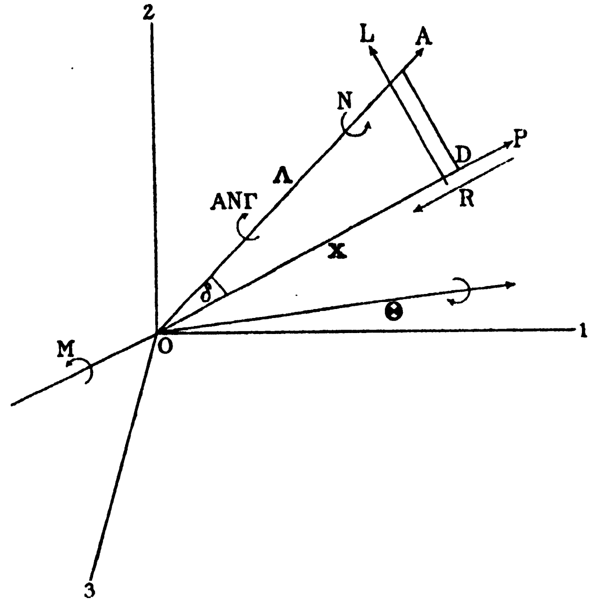


Fig. 9.

With the notation already introduced in Part I., the total angular momentum of the shell can be expressed as the sum of two vectors :—

(i.) The angular momentum about OA, $AN\boldsymbol{\Lambda}$;

(ii.) The total angular momentum about a transverse axis.

If the total angular velocity about a transverse axis is ω, the angular momentum is $B\omega$, and is equal to the moment of momentum of a particle whose mass is B and whose distance from O is represented by the vector $\boldsymbol{\Lambda}$. Now the actual velocity of such a particle relative to O is $\boldsymbol{\Lambda}'-[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Theta}]$, and therefore its moment of momentum about O is

$$B\{[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Lambda}']-[\boldsymbol{\Lambda}\,.[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Theta}]]\}.$$

The total angular momentum, $\mathbf{H}$, of the shell about O is therefore given by the equation

$$\mathbf{H} = AN\boldsymbol{\Lambda}+B\{[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Lambda}']-[\boldsymbol{\Lambda}\,.[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Theta}]]\}\,; \tag{3.101}$$

using (3.012) this becomes

$$\mathbf{H} = AN\boldsymbol{\Lambda}+B\{[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Lambda}']-(\boldsymbol{\Lambda}\,.\,\boldsymbol{\Theta})\,\boldsymbol{\Lambda}+\boldsymbol{\Theta}\}. \tag{3.102}$$

---

* *I.e.* $(\boldsymbol{\Lambda})^2 = (\mathbf{X})^2 = 1$.

The force components that we propose to include at this stage are R, L, M, and $\mathrm{AN}\Gamma$. To simplify the algebra we write*

$$\mathrm{L} = \kappa m v \sin \delta, \qquad \mathrm{M} = \mu \sin \delta.$$

The various components can then be represented by the following vectors:—

(i.) The drag R, by the vector $-\mathrm{R}\mathbf{\Lambda}$;

(ii.) The cross-wind force L, by the vector† $\kappa m v \{\mathbf{\Lambda} - \mathbf{X} \cos \delta\}$;

(iii.) The couple M, by the vector† $\mu [\mathbf{X} . \mathbf{\Lambda}]$;

(iv.) The couple $\mathrm{AN}\Gamma$, by the vector $-\mathrm{AN}\Gamma\mathbf{\Lambda}$.

The complete equation for the angular motion is therefore

$$\mathbf{H}' - [\mathbf{H} . \mathbf{\Theta}] = \mu [\mathbf{X} . \mathbf{\Lambda}] - \mathrm{AN}\Gamma\mathbf{\Lambda}, \tag{3.103}$$

where $\mathbf{H}$ is given by (3.102). Taking the scalar product of both sides of (3.103) into $\mathbf{\Lambda}$, we obtain, with the help of (3.011)-(3.013),

$$\mathrm{N}' = -\mathrm{N}\Gamma. \tag{3.104}$$

After substituting for N′, equation (3.103), written in full, reduces to

$$\begin{aligned} \mathrm{AN}\mathbf{\Lambda}' + \mathrm{B}[\mathbf{\Lambda} . \mathbf{\Lambda}''] - 2\mathrm{B}(\mathbf{\Lambda} . \mathbf{\Theta})\mathbf{\Lambda}' - \mathrm{B}(\mathbf{\Lambda} . \mathbf{\Theta}')\mathbf{\Lambda} + \mathrm{B}\mathbf{\Theta}' \\ - \mathrm{AN}[\mathbf{\Lambda} . \mathbf{\Theta}] + \mathrm{B}(\mathbf{\Lambda} . \mathbf{\Theta})[\mathbf{\Lambda} . \mathbf{\Theta}] = \mu [\mathbf{X} . \mathbf{\Lambda}]. \end{aligned} \tag{3.105}$$

3.11. *The Equations of Motion of the Centre of Gravity.*—The velocity of the centre of gravity is represented by the vector $v\mathbf{X}$, and its acceleration is therefore represented by the vector

$$\frac{d}{dt}\{v\mathbf{X}\} - v[\mathbf{X} . \mathbf{\Theta}].$$

In addition to the drag and cross-wind force impressed by the air, we shall suppose that gravity is acting on the shell.

---

* The mass and velocity of the shell are m and v respectively. For the rest of the notation see § 1.31.

† If a perpendicular AD be drawn from A to OP, DA is parallel to the direction of the cross-wind force L, and its length is sin δ, if OA is of unity length. The vector DA is equal to the difference of the vectors OA and OD, so that it is equal to $\mathbf{\Lambda} - \mathbf{X} \cos \delta$. Hence $\{\mathbf{\Lambda} - \mathbf{X} \cos \delta\}/\sin \delta$ is the unit vector parallel to the cross-wind force. Similarly $[\mathbf{X} . \mathbf{\Lambda}]/\sin \delta$ is the unit vector normal to the plane AOP *i.e.*, parallel to the axis of the couple M. It is easy to verify, with the help of (3.012), that

$$\mathbf{\Lambda} - \mathbf{X} \cos \delta = [[\mathbf{X} . \mathbf{\Lambda}] . \mathbf{X}].$$

The acceleration due to gravity is represented by the vector $\mathbf{G}$, whose modulus is g.[*] Under these conditions the vector equation of motion of the centre of gravity is

$$(3.111)\qquad \frac{d}{dt}\{v\mathbf{X}\}-v[\mathbf{X}\,.\,\boldsymbol{\Theta}] = -\frac{\mathrm{R}}{m}\mathbf{X}+\kappa v\{\boldsymbol{\Lambda}-\mathbf{X}\cos\delta\}+\mathbf{G}.$$

Taking the scalar product of both sides into $\mathbf{X}$, equation (3.111) reduces to

$$(3.112)\qquad v' = -\mathrm{R}/m+(\mathbf{G}\,.\,\mathbf{X}).$$

On substituting this value of v' in (3.111), and dividing by v, we obtain

$$(3.113)\qquad \mathbf{X}'-[\mathbf{X}\,.\,\boldsymbol{\Theta}] = \kappa\{\boldsymbol{\Lambda}-\mathbf{X}\cos\delta\}+\{\mathbf{G}-(\mathbf{G}\,.\,\mathbf{X})\mathbf{X}\}/v.$$

Equations (3.104), (3.105), (3.112), and (3.113) determine the motion completely.

§ 3.2. *Equations of Motion of Type* α.

When a shell is initially sufficiently stable, and leaves the muzzle so that its initial disturbance is small, it will be shown[†] that the axis OA and the direction of motion OP deviate, at any time t, by small angles only from the direction of the tangent to the corresponding[‡] plane trajectory at the same time. This is true of the early part of all trajectories, and for the whole of a trajectory whose initial elevation is less than 45 degrees—at any rate, when the muzzle velocity is fairly large Under these circumstances we may follow the classical[§] treatment in regarding the plane trajectory as a first approximation to the actual trajectory. It is then convenient to refer the motion to axes moving with the tangent to this plane trajectory. The axis O1 is the tangent to the plane trajectory drawn in the direction of motion; axis O2 is the upward normal; and axis O3 is horizontal and to the right, as viewed from the gun. The components of $\boldsymbol{\Lambda}$ and $\mathbf{X}$ are (l, m, n) and (x, y, z), which are therefore the direction cosines of OA and OP respectively.

It will now be shown that it is possible to express the complete motion approximately in terms of the two complex variables, $m+in$ and $y+iz$, and the elements of the plane trajectory. We suppose that the equations of the plane trajectory have been numerically solved, so that, *e.g.*, v_1 and θ_1, the velocity and inclination in the plane trajectory, may be regarded as tabulated functions of t.

---

[*] The vector $\mathbf{G}$ may, if desired, be regarded as representing any force which acts through the centre of gravity and is a function of position only.

[†] See § 4.21.

[‡] The *corresponding* plane trajectory is the trajectory which would be described by the same shell, with the same initial velocity and initial direction of motion, if its yaw remained always zero.

See, *e.g.*, CRANZ, 'Zeitschrift für Math. u. Phys.' The equations we obtain, however, appear to be new.

The components of **Θ** are $(0, 0, \theta'_1)$. Using the foregoing values of the components of **Λ**, **X** and **Θ** in equations (3.104) and (3.105), we obtain

$$\text{(3.201)} \qquad N' = -N\Gamma,$$

as before; the second and third components of (3.105) give

$$\text{(3.202)} \qquad ANm' + B(nl'' - ln'') - 2Bnm'\theta'_1 - Bmn\theta''_1 + ANl\theta'_1 - Bnl\theta'^2_1 = \mu(zl - xn),$$

$$\text{(3.203)} \qquad ANn' + B(lm'' - ml'') - 2Bnn'\theta'_1 - Bn^2\theta''_1 + B\theta''_1 = \mu(xm - yl).$$

To solve the equations it is necessary to neglect certain terms. A discussion of the relative magnitude of the terms neglected, in various circumstances, will be given later in § 4.3. Some of these terms are negligible in all cases, on account of the smallness of θ'_1 in comparison with the angular spin of the shell. Others are only negligible so long as the yaw δ is so small that $1-\cos\delta$ and $1-\sin\delta/\delta$ may be neglected in comparison with unity. By such arguments it is not difficult to justify the reduction of these equations to the form

$$\text{(3.204)} \qquad ANm' - Bn'' + AN\theta'_1 = \mu(z - n),$$

$$\text{(3.205)} \qquad ANn' + Bm'' + B\theta''_1 = \mu(m - y).$$

For the particular case of the initial motion of the shells from the gun rifled 1 turn in 30 calibres in the present trials, the terms neglected are, in general, less than 1 per cent. of some term retained, and the coefficients of equations (3.204) and (3.205) may be regarded as affected by possible 1 per cent. errors. Even in the case of the gun rifled 1 turn in 40 calibres, where values of δ as great as 7 degrees or more are met with among the stable rounds, the employment of (3.204) and (3.205) is justifiable.

We now define new variables and constants by the equations

$$\eta + c\xi = m + in, \qquad c\xi = y + iz,$$
$$AN/B = \Omega, \qquad c = \cos\theta_1, \qquad \theta'_1 + i\theta''_1/\Omega = \Phi.$$

If we multiply (3.204) by i, and subtract from (3.205), we obtain

$$\text{(3.206)} \qquad \frac{d^2}{dt^2}(\eta + c\xi) - i\Omega\frac{d}{dt}(\eta + c\xi) - \frac{\mu\eta}{B} = i\Omega\Phi.$$

So long as the yaw remains small, equations (3.201) and (3.206) may be taken as equivalent to (3.104) and (3.105).

3.21. *The Motion of the Centre of Gravity.*—With the present axes, the components of **G** are $(-g\sin\theta_1, -g\cos\theta_1, 0)$. Equation (3.112) becomes†

$$\text{(3.211)} \qquad v' = -R(v, \delta)/m^* - g(x\sin\theta_1 + y\cos\theta_1).$$

---

† To avoid confusion the mass of the shell is temporarily denoted by m^*

The second and third components of (3.113) become

$$(3.212) \qquad y' + x\theta_1' = \kappa(m - y\cos\delta) - (g/v)\cos\theta_1 + (yg/v)(x\sin\theta_1 + y\cos\theta_1),$$

$$(3.213) \qquad z' = \kappa(n - z\cos\delta) + (zg/v)(x\sin\theta_1 + y\cos\theta_1).$$

The equation of the plane trajectory corresponding to (3.211) ($\delta = y = 0$, $x = 1$) is

$$(3.214) \qquad v'_1 = -\mathrm{R}(v_1, 0)/m^* - g\sin\theta_1.$$

Therefore, if $u = v - v_1$, u satisfies the equation

$$(3.2141) \qquad u' = -\{\mathrm{R}(v_1 + u, \delta) - \mathrm{R}(v_1, 0)\}/m^* - g\{(x-1)\sin\theta_1 + y\cos\theta_1\}.$$

In §§ 4.22, 4.31, we shall show that it is legitimate to regard the value of u determined by this equation as zero. We can therefore replace v by v_1 in (3.206), (3.212) and (3.213).

A further discussion shows that (3.212) and (3.213) can be reduced to

$$y' = \kappa(m - y) + (g/v_1)\, y\sin\theta_1,$$
$$z' = \kappa(n - z) + (g/v_1)\, z\sin\theta_1,$$

the accuracy and validity of these equations being the same as those of (3.204) and (3.205).[†] These equations combine to give

$$\frac{d}{dt}(c\zeta) = \kappa\eta - \frac{gc\sin\theta_1}{v_1}\zeta;$$

or, using the equation of the plane trajectory, $\theta'_1 = -(g/v_1)\cos\theta_1$,

$$(3.215) \qquad \zeta' = \kappa\eta/c.$$

In the cases contemplated this equation is equivalent to (3.212) and (3.213). Then (3.215), (3.206) and the equations of the plane trajectory represent the required approximation to the complete equations of motion of the shell.

In order to convert (3.206) and (3.215) into linear differential equations, it is necessary to assume that μ and κ are independent of δ, and regard them as functions of v_1. This approximation involves errors no greater than the previous approximations. If Ω is treated as a variable, it must be determined by (3.201), Γ being regarded as a known function of the time. All the coefficients in (3.215) and (3.206) are then known functions of the time.

§ 3.3. *Equations of Motion of Type β.*

In the neighbourhood of the vertex of a trajectory of elevation as great as 70 degrees, the yaw, as stated in § 1.34, may reach large values. In such cases, the plane trajectory can no longer be regarded as a valid first approximation, and the only possible method is to obtain equations of motion which are suitable for direct step-by-step integration. For this purpose the following set of moving axes are most suitable, as they reduce the equations of motion of the centre of gravity to its simplest form. We take the true direction of motion OP for the axis 1 and a horizontal line at right angles to OP for the axis 3. We define the position of OP by spherical polar co-ordinates θ, ψ with respect to axes fixed in direction at O, see fig. 10. Then $\mathbf{X}$ has components $(1,\ 0,\ 0)$, $\boldsymbol{\Theta}$ has components $(-\psi' \sin\theta,\ -\psi' \cos\theta,\ \theta')$, $\mathbf{G}$ has components $(-g\sin\theta,\ -g\cos\theta,\ 0)$ and $\boldsymbol{\Lambda}$ components $(l,\ m,\ n)$ as before.

[†] With the exception noted in § 4.22.

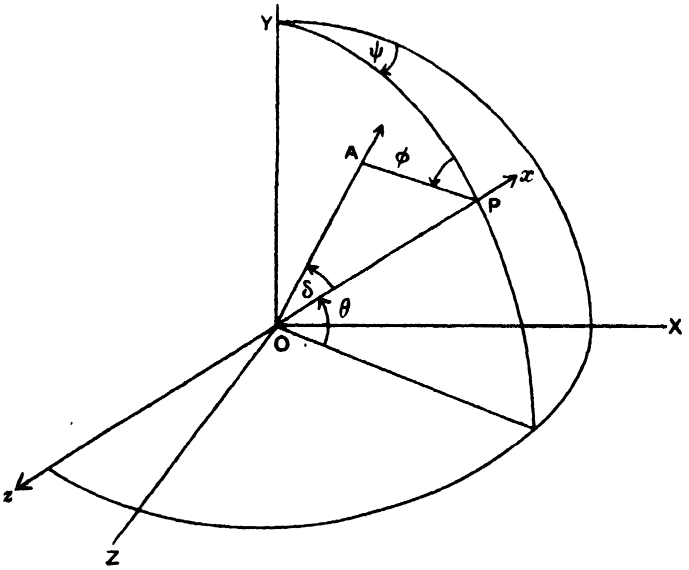


Fig. 10. OX, Y, Z are fixed axes, OY being the upwards vertical; the plane XOY contains the line of fire.

Equation (3.105), when written out in full, becomes very complicated. To simplify it, we can, under certain circumstances, neglect the angular momentum about a transverse axis compared to the angular momentum about the axis of the shell. The legitimacy of this approximation, which is equivalent to putting $B = 0$ in (3.105), is discussed in § 4.33. It should be stated that this type of approximation also is classical,* but that the equations we obtain are apparently new and of a wide range of validity.

As before, we have

$$N' = -N\Gamma. \tag{3.301}$$

The second and third components of (3.105) reduce to

$$\begin{aligned} ANm' - AN(-n\psi' \sin\theta - l\theta') &= -\mu n, \\ ANn' - AN(-l\psi' \cos\theta + m\psi' \sin\theta) &= \mu m, \end{aligned}$$

---

* See, *e.g.*, CHARBONNIER, 'Traité de Balistique Extérieure,' Livre V., Chap. IV.

or, writing ω for μ/AN,

$$m' = -n\,(\omega+\psi'\sin\theta)-l\theta', \tag{3.302}$$

$$n' = m\,(\omega+\psi'\sin\theta)-l\psi'\cos\theta. \tag{3.303}$$

The corresponding equations of motion of the centre of gravity are

$$v' = -\mathrm{R}\,(v,\,\delta)/m^{*}-g\sin\theta, \tag{3.304}$$

$$\theta' = \kappa m-(g/v)\cos\theta, \tag{3.305}$$

$$\psi'\cos\theta = \kappa n. \tag{3.306}$$

The six equations (3.301) to (3.306) can be solved by a step-by-step process, if R, κ, μ and Γ are numerically known functions of v and δ. They are valid without restriction as to the size of δ, and have proved of value for the discussion of trajectories at very high elevations. They are, however, necessarily invalid when any question of stability is under discussion.

§ 3.4. *Equations of Motion of Type* γ.

For the purpose of discussing the initial motion of a shell which is unstable or just stable, equations of types α and β are invalid, and it is necessary to make use of equations corresponding to the equations of energy and angular momentum for a top. The equations we shall thus obtain are of far less general applicability than types α and β.

With this object we take the scalar product of both sides of equation (3.105) into the vector $[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Lambda}']+\boldsymbol{\Theta}$, and obtain, after reduction,

$$\tfrac{1}{2}\mathrm{B}\frac{d}{dt}\left\{(\boldsymbol{\Lambda}')^2+2\,(\boldsymbol{\Theta}\,.\,[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Lambda}'])+(\boldsymbol{\Theta})^2-(\boldsymbol{\Lambda}\,.\,\boldsymbol{\Theta})^2\right\} = -\mu\,(\mathbf{X}\,.\,\{\boldsymbol{\Lambda}'-[\boldsymbol{\Lambda}\,.\,\boldsymbol{\Theta}]\}). \tag{3.401}$$

Using the axes described in the last section, we note that, over a limited range at the beginning of a trajectory, the first two components of $\boldsymbol{\Theta}$ are numerically very small compared to the third, θ'. We shall find that the effect of θ' itself is negligible in the cases we consider. We shall therefore neglect the other components of $\boldsymbol{\Theta}$ at once. Taking δ and ϕ as spherical polar co-ordinates of the axis OA referred to the moving axes, so that

$$l = \cos\delta,\ m = \sin\delta\cos\phi,\ n = \sin\delta\sin\phi,\dagger$$

---

† The angle ϕ is not exactly the angle measured by the jump cards, but the difference is negligible. The angle δ is exactly the measured angle of yaw.

we find that (3.401) reduces to

$$(3.402)\quad \tfrac{1}{2}\mathrm{B}\frac{d}{dt}\left\{\delta'^2+\phi'^2\sin^2\delta+2\theta'(\delta'\cos\phi-\phi'\sin\delta\cos\delta\sin\phi)+\theta'^2(1-\sin^2\delta\sin^2\phi)\right\}$$
$$=-\mu\left\{\frac{d\cos\delta}{dt}-\theta'\sin\delta\cos\phi\right\}.$$

This is the equation corresponding to the equation of energy. The first component of (3.105) corresponds to the equation of angular momentum for a top, and reduces in the same way to

$$(3.403)\quad \frac{d}{dt}\{\mathrm{AN}\cos\delta+\mathrm{B}\phi'\sin^2\delta\}+2\mathrm{B}\theta'\delta'\sin^2\delta\sin\phi-\mathrm{B}\theta''\sin\delta\cos\delta\sin\phi$$
$$-\mathrm{AN}\theta'\sin\delta\cos\phi+\mathrm{B}\theta'^2\sin^2\delta\cos\phi\sin\phi=0.$$

Equation (3.201) remains unaltered. Over the range of the jump card experiments a mean value of δ'/θ' is 50. We shall therefore regard it as legitimate for our present purposes to neglect all terms containing θ'. On integrating the resulting equations we obtain

$$(3.404)\qquad \tfrac{1}{2}\mathrm{B}(\delta'^2+\phi'^2\sin^2\delta)+\int_0^\delta\mu d\cos\delta=\tfrac{1}{2}\mathrm{BE},$$

$$(3.405)\qquad \mathrm{AN}\cos\delta+\mathrm{B}\phi'\sin^2\delta=\mathrm{BF},$$

where E and F are constants of integration. In (3.405) it is assumed that N is constant. If μ is constant these equations are of the same form as those of the motion of a top. In the more important applications to the jump card trial which we shall make of (3.404) and (3.405), μ will be treated as a variable function of δ, and also of v.

§ 3.5. *The Additional Force Components* H, J *and* K.

It is now necessary to consider the effect of the additional force components, mentioned in §§ 1.1, 1.12, and denoted by H, J and K. These have so far been omitted from the general equations for the sake of simplicity. The couples H and J will affect the angular motion of the axis, and the force K will affect the motion of the centre of gravity. For algebraic convenience we define new variables h, γ, λ, by the equations

$$\mathrm{H}=h\mathrm{B}w,\qquad \mathrm{J}=\mathrm{AN}\gamma\sin\delta,\qquad \mathrm{K}=m\mathrm{N}v\lambda\sin\delta,$$

where w is the total angular velocity of the axis of the shell. The force components may then be represented, in the notation of § 3.1, by the following vectors:—

H by the vector $-h\mathrm{B}\{[\mathbf{\Lambda}\,.\,\mathbf{\Lambda}']-(\mathbf{\Lambda}\,.\,\mathbf{\Theta})\,\mathbf{\Lambda}+\mathbf{\Theta}\}$;

J by the vector $\mathrm{AN}\gamma(\mathbf{\Lambda}\cos\delta-\mathbf{X})$;

K by the vector $m\mathrm{N}v\lambda\,[\mathbf{\Lambda}\,.\,\mathbf{X}]$.

To include the effect of these components we add to the right-hand side of (3.105)

$$-h\mathrm{B}\{[\mathbf{\Lambda}\,.\,\mathbf{\Lambda}']-(\mathbf{\Lambda}\,.\,\mathbf{\Theta})\,\mathbf{\Lambda}+\mathbf{\Theta}\}+\mathrm{AN}\gamma\,(\mathbf{\Lambda}\cos\delta-\mathbf{X}),$$

and to the right right-hand side of (3.113)

$$\mathrm{N}\lambda\,[\mathbf{\Lambda}\,.\,\mathbf{X}].$$

Equations (3.104) and (3.112) (type α) are unaltered. As a result, the following additions must be made to the right-hand side of succeeding equations :—

To (3.202), $+h\mathrm{B}n'+\mathrm{AN}\gamma\,(m-y)$.
To (3.203), $-h\mathrm{B}m'+\mathrm{AN}\gamma\,(n-z)$.
To (3.212), $-\mathrm{N}\lambda\,(z-n)$.
To (3.213), $-\mathrm{N}\lambda\,(m-y)$.

As the total effect of the extra components h, γ and λ is certainly small in any practical case, we have neglected all terms other than those of the lowest order in δ. Equations (3.206) (3.215), when modified by the inclusion of these extra terms, become

$$\frac{d^2}{dt^2}(\eta+c\zeta)-(i\Omega-h)\frac{d}{dt}(\eta+c\zeta)-\left(\frac{\mu}{\mathrm{B}}-i\Omega\gamma\right)\eta=i\Omega\Phi\,; \tag{3.501}$$

$$\zeta'=(\kappa-i\mathrm{N}\lambda)\,\eta/c. \tag{3.502}$$

3.51. *The Additional Terms in Equations of Types β and γ.*—The additional terms in the equations of type β can be written down in a similar manner. The following additions must be made to the right-hand sides of the equations :—

To (3.302), $+\gamma m\cos\delta-h\,(nl'-ln')/\Omega$.
To (3.303), $+\gamma n\cos\delta-h\,(lm'-ml')/\Omega$.
To (3.305), $+\mathrm{N}\lambda n$.
To (3.306), $-\mathrm{N}\lambda m$.

The terms in h are negligible, as they are $\mathrm{O}\,(h\delta'/\Omega^2)$ compared with the principal terms $-n\omega-l\theta'$, so long as ω/Ω is not very small. The principal application of these equations is to the motion of a shell near the vertex of a trajectory at an elevation of 70 degrees, where the velocity becomes small while the spin probably remains large. Under these circumstances the terms γ and λ arising from the spin rise in importance relatively to the terms ω and κ representing the ordinary force components. The inclusion of the extra terms γ and λ in these equations is at present of no practical importance, as we have no definite information as to their value.

The corresponding terms could be added to equations of type γ by the same methods, but the results are of no importance as it is impossible at present to solve these equations unless these terms are neglected.

§ 3.6. *The Approximate Solution of Equations of Type* α.

3.61. *The Nature of the Solution Required.*—The system of equations (3.501), (3.502) are linear differential equations with respect to the time of the second order, the coefficients being regarded as known functions of the time t. Since these functions are in practice empirical and by no means simple, an exact solution is impossible. To simplify the discussion we write

$$s = \mathrm{A}^2\mathrm{N}^2/4\mathrm{B}\mu, \qquad \mu/\mathrm{B} = \Omega^2/4s,$$

so that equations (3.501), (3.502) become

$$\frac{d^2}{dt^2}(\eta+c\zeta)-(i\Omega-h)\frac{d}{dt}(\eta+c\zeta)-\left(\frac{\Omega^2}{4s}-i\Omega\gamma\right)\eta = i\Omega\Phi; \tag{3.611}$$

$$\zeta'-(\kappa-i\mathrm{N}\lambda)\,\eta/c = 0. \tag{3.612}$$

If the terms in ζ, h, γ be omitted from (3.611), and s, N and Ω are assumed constant, the equation reduces to that for the small oscillations of a top in the neighbourhood of the vertical.

The coefficient s is the stability factor as defined in § 1.31. In order to be able to apply the approximations on which (3.611) and (3.612) are based, we shall find that it is necessary to assume that the shell is more than just stable, *e.g.*, $s > 1{\cdot}1$.

We proceed to develop an approximate solution of the equations on the assumption that Ω is large. If we ignore the dimensions of the various terms, and take the unit of time as 1 second, then Ω is in practice greater than 100 (radians per second), all other terms being of the order unity. This is really equivalent to assuming that all the ratios κ/Ω, h/Ω, ..., which are of no dimensions, are small. It will be found necessary to assume further that all derivatives with respect to the time are of order unity in units of 1 second, *e.g.*, that κ', κ'', s', Ω'..., are of order unity. These conditions are satisfied in practice. As a result, we can say that κ/Ω, s'/Ω..., are small quantities of the first order, and κ'/Ω^2, s''/Ω^2, θ''_1/Ω^2..., are small quantities of the second order. For simplicity, we shall throughout ignore dimensions, and denote such terms of the first order by $O(1/\Omega)$, and terms of the second order by $O(1/\Omega^2)$.[*] The arithmetical values of the various terms are investigated in detail in § 4.3 below.

The above facts indicate the lines on which an approximate solution is to be sought—*we require the asymptotic expansion of the solution (or its leading terms) for large values of the parameter* Ω. Methods of obtaining such expansions have been investigated in general terms by HORN and SCHLESINGER.* A method, which is slightly different algebraically, is more convenient here; the asymptotic properties of our solutions, however, may be regarded as established by the researches of these authors.

---

[*] In practice the spin N, and therefore Ω, decreases slightly along the trajectory, but the diminution is not sufficient to affect the assumption that Ω is large.

The equations (3.611) and (3.612) are a pair of linear differential equations with respect to the time for the two dependent variables η and ζ, (3.611) being of the second order and (3.612) of the first. There must, therefore, be three independent solutions.

It is convenient to eliminate ζ' and ζ'' from (3.611) by the use of (3.612), the result being

$$(3.613) \qquad \eta''-(i\Omega-h-\kappa_1)\,\eta'-\{\Omega^2/4s+i\Omega\,(\kappa_1-\gamma)-h\kappa_1-\kappa'_1-\kappa_1c'/c\}\,\eta \\ -\{i\Omega c'-hc'-c''\}\,\zeta = i\Omega\Phi,$$

where κ_1 is written for $(\kappa-i\mathrm{N}\lambda)$. It is believed that $\mathrm{N}\lambda$ is small compared with κ, so that for simplicity the term $\mathrm{N}\lambda$ will usually be omitted in subsequent work. The term γ will however be retained.

3.62. *The Complementary Function.*—A first approximation to the three independent complementary functions is obtained, following HORN and SCHLESINGER, by making the substitution,

$$\eta = e^{i\Omega x}\bar{\eta}, \qquad \zeta = e^{i\Omega x}\zeta,$$

and treating $\bar{\eta}$ and $\bar{\zeta}$ as constants in determining η' and ζ'. We also neglect all but the highest order terms in Ω in each equation. The equations then reduce to

$$(3.621) \qquad (-\Omega^2x'^2+\Omega^2x'-\Omega^2/4s)\,\bar{\eta}-i\Omega c'\bar{\zeta} = 0;$$

$$(3.622) \qquad -\kappa\bar{\eta}/c+i\Omega x'\bar{\zeta} = 0.$$

On eliminating $\bar{\eta}$ and ζ, and retaining only the terms of highest order in Ω, these reduce to

$$x'\,(x'^2-x'+1/4s) = 0,$$

a cubic equation for x' whose three roots correspond to the three independent solutions required. The roots are $\frac{1}{2}\pm\frac{1}{2}(1-1/s)^{\frac{1}{2}}$, 0, or writing, for shortness, $\sigma = (1-1/s)^{\frac{1}{2}}$, the three values of x are

---

* J. HORN, 'Mathematische Annalen,' vol. 52, p. 271 and p. 340. L. SCHLESINGER, *ibid.*, vol. 63, p. 277; 'Comp. Rend.,' vol. 142, p. 1031. The investigations of the complementary function given by these writers are fairly complete, the asymptotic nature of the expansions being established. The latter writer considers a system of n linear differential equations. A similar treatment of the complementary function and the particular integral of a special equation is suggested (without proof) by M. DE SPARRE 'Atti (Rendiconti) della R. Acc. dei Lincei,' 1898, Ser. V., vol. 72, p. 111; this writer was obviously led to the solution he gives by his researches on the motion of spinning projectiles.

[*Note added July* 30, 1920. See also G. D. BIRKHOFF, 'Trans. Amer. Math. Soc', vol. 9, p. 219.]

$$x_1 = \tfrac{1}{2}\int_0^t (1+\sigma)\,dt\,; \qquad x_2 = \tfrac{1}{2}\int_0^t (1-\sigma)\,dt\,; \qquad x_3 = 0.$$

It appears that the first two solutions correspond to the complementary function of equation (3.613) with the term in ζ neglected, so that η is large compared with ζ. If $s < 1$, σ is imaginary, the motion is unstable and the solution fails. In the third solution ζ is large compared with η, and a first approximation to it gives a constant value to ζ, obtained by neglecting the term in η in equation (3.612). It is convenient to obtain the first two solutions independently by a special method.

We first omit the term in ζ in equation (3.613); it is not required till the second approximation. Write the equation, for simplicity, in the form

(3.623) . $$\eta''-i\mathrm{A}\eta'-\mathrm{B}\eta = 0,$$

where

$$\mathrm{A} = \Omega+ih+i\kappa,$$
$$\mathrm{B} = \Omega^2/4s+i\Omega(\kappa-\gamma)-h\kappa-\kappa'-\kappa c'/c.$$

Remove the second term by substituting

$$\eta = y\exp\left\{\tfrac{1}{2}i\int_0^t \mathrm{A}\,dt\right\},$$

giving

(3.6231) $$y''+\mathrm{M}y = 0,$$

where

$$\mathrm{M} = \tfrac{1}{4}\mathrm{A}^2-\mathrm{B}+\tfrac{1}{2}i\mathrm{A}'$$
$$= \tfrac{1}{4}\Omega^2\sigma^2\left\{1+\frac{2i}{\Omega\sigma^2}(h-\kappa+2\gamma+\mathrm{N}'/\mathrm{N})+\mathrm{O}\left(\frac{1}{\Omega^2}\right)\right\}.$$

Substitute $y = \mathrm{R}e^{\pm i\mathrm{P}}$, so that (3.6231) becomes

(3.6232) $$\mathrm{R}''\pm(2i\mathrm{P}'\mathrm{R}'+i\mathrm{P}''\mathrm{R})-\mathrm{P}'^2\mathrm{R}+\mathrm{MR} = 0.$$

We may make P and R satisfy any single relation we choose, *e.g.*,

$$2\mathrm{P}'\mathrm{R}'+\mathrm{P}''\mathrm{R} = 0,$$

giving $\mathrm{P}' = 1/\mathrm{R}^2$,* so that (3.6232) becomes

(3.6233) . $$\mathrm{R}''-1/\mathrm{R}^3+\mathrm{MR} = 0.$$

---

* More generally $\mathrm{P}' = a/\mathrm{R}^2$, where a is a constant, but the value of this constant is immaterial, as it disappears in the result.

This equation may be solved asymptotically, by successive approximation, by writing*

$$R = R_0(1+R_1+R_2 \ldots),$$

where

$$R_1 = O(1/M) = O(1/\Omega^2), \qquad R_2 = O(1/\Omega^4) \ldots .$$

We obtain, in succession, the approximations

$$R_0 = M^{-\frac{1}{4}},$$

$$R_1 = -\tfrac{1}{4}M^{-\frac{3}{4}}\frac{d^2}{dt^2}(M^{-\frac{1}{4}}),$$

verifying the relation $R_1 = O(1/\Omega^2)$. The order of magnitude of R_1 in practice will be discussed in § 4.32, where it will be shown to be negligible.† We therefore take as our two standard solutions

$$y_1 = M^{-\frac{1}{4}} \exp\left\{i\int_0^t M^{\frac{1}{2}}\,dt\right\},$$

$$y_2 = M^{-\frac{1}{4}} \exp\left\{-i\int_0^t M^{\frac{1}{2}}\,dt\right\},$$

giving, for the complementary function of (3.623),

(3.6234) $$\eta = (\Omega\sigma)^{-\frac{1}{2}}\{1+O(1/\Omega)\}\{K_1e^{iP_1}+K_2e^{iP_2}\},$$

where K_1, K_2 are arbitrary constants, and P_1, P_2 are given by

$$P_1,\ P_2 = \tfrac{1}{2}\int_0^t [\Omega(1\pm\sigma)+i\{h+\kappa\pm(h-\kappa+2\gamma+N'/N)/\sigma\}]\,dt.$$

This is the form of solution which is used in analysing the jump card experiments, and contains all the terms that can be required in practice.

It is now necessary to examine the effect of the term in ζ in (3.613), which has so far been omitted. The value of ζ', obtained from (3.612), corresponding to the first solution for η, is

$$\zeta' = (\kappa/c)(\Omega\sigma)^{-\frac{1}{2}}e^{iP_1},$$

so that, on integrating by parts to obtain the leading terms,

$$\zeta_1 = \frac{2\kappa\eta_1}{ic\Omega(1+\sigma)} + O\left(\frac{1}{\Omega^2}\right).$$

---

* At this point the advantage of our *ad hoc* method over more general methods is apparent, as we obtain in one step a solution with an error $O(1/\Omega^2)$, whereas the general method requires two steps.

† We assume that the numerical value of R_1, the next term in the expansion, is a measure of the error in the solution caused by omitting all terms after the first. The expansions for y_1 and y_2 are known to be asymptotic for large values of Ω, so that the error will be some finite multiple of R_1, but the size of the numerical factor is unknown.

Similarly we have

$$\zeta_2 = \frac{2\kappa\eta_2}{ic\Omega\,(1-\sigma)} + \mathrm{O}\left(\frac{1}{\Omega^2}\right),$$

verifying that ζ is small compared to η. The contribution of ζ to equation (3.613) is thus

(3.6235)
$$-\frac{2\kappa c'\eta}{c\,(1\pm\sigma)} + \mathrm{O}\left(\frac{1}{\Omega}\right),$$

which is equivalent to an addition to the coefficient of η of terms which are $\mathrm{O}\,(1/\Omega^2)$ compared to the principal term. The solution can be repeated with these terms included, but is unaffected to the order to which we are working. We shall take our first two standard solutions in the form

(3.624)
$$\eta_1 = \left(\frac{\Omega_0\sigma_0}{\Omega\sigma}\right)^{\frac{1}{2}} e^{i\mathrm{P}_1}, \qquad \zeta_1 = \frac{2\kappa\eta_1}{ic\Omega\,(1+\sigma)},$$

(3.625)
$$\eta_2 = \left(\frac{\Omega_0\sigma_0}{\Omega\sigma}\right)^{\frac{1}{2}} e^{i\mathrm{P}_2}, \qquad \zeta_2 = \frac{2\kappa\eta_2}{ic\Omega\,(1-\sigma)}.$$

The differential coefficients of the solutions may be obtained by differentiation of these equations.

For the third solution we have shown that the exponential index is zero to the first order, and that a first approximation is given by

$$\eta_3 = \zeta'_3 = 0, \qquad \zeta_3 = 1.$$

The expansions take a somewhat different form, like those for the particular integral, and we write

(3.6261)
$$\eta_3 = \eta^{(0)} + \eta^{(1)}/i\Omega + \eta^{(2)}/(i\Omega)^2 \ldots,$$

(3.6262)
$$\zeta_3 = \zeta^{(0)} + \zeta^{(1)}/i\Omega + \zeta^{(2)}/(i\Omega)^2 \ldots.$$

Substituting in equations (3.612) and (3.613), we obtain

$$\eta^{(1)} = 4sc' = -4s\,\theta'_1 \sin\theta_1,$$

$$\zeta^{(1)} = 4\int_0^t \kappa sc'dt/c.$$

The significance of this solution will be considered after the particular integral has been discussed.

Our standard third solution is then

(3.627)
$$\eta_3 = \frac{4sc'}{i\Omega}, \qquad \zeta_3 = 1 + \frac{4}{i\Omega}\int_0^t \frac{\kappa sc'}{c}\,dt.$$

3.63. *The Particular Integral.*—The question of the particular integral is not treated generally by HORN and SCHLESINGER. The former considers shortly a very particular case.* Their methods can, however, be extended to obtain the results we require.

We assume an expansion for the particular integral $\bar{\eta}$, $\bar{\zeta}$ of the form of (3.6261), (3.6262). This integral can be specified in such a way that initially $\bar{\zeta} = 0$, *i.e.*, $\zeta^{(0)} = \zeta^{(1)} = \ldots = 0$, and $\eta^{(0)} = 0$. It will then be found to be unique.

On substituting in equations (3.612) and (3.613), with the right-hand side retained, and equating powers of $i\Omega$, we find first that $\zeta^{(0)} = \eta^{(0)} = 0$ for all time, and then

$$(3.631) \qquad \eta^{(1)} = 4s\Phi, \qquad \zeta^{(1)} = \int_0^t 4s\kappa\Phi dt/c.$$

The first two terms in the expansions of $\bar{\eta}$ and $\bar{\zeta}$ take the forms

$$(3.632) \qquad \bar{\eta} = \frac{4s\Phi}{i\Omega} + \frac{4s}{(i\Omega)^2}\left\{\frac{d}{dt}(4s\Phi) + (\kappa-\gamma)\,4s\Phi + c'\int_0^t 4s\kappa\Phi dt/c\right\};$$

$$(3.633) \qquad \bar{\zeta} = \int_0^t \kappa\bar{\eta}dt/c.$$

Equations (3.632) and (3.633) will be taken as the standard particular integral. Since, moreover, they contain no periodic terms, and the initial value of $\bar{\zeta}$ is zero and those of $\bar{\eta}$ and $\bar{\eta}'$ very small in practice, it is convenient to take this solution as the standard solution of the equations of motion in cases where the initial values of η and η' are not exactly known—*e.g.*, in calculating the drift.

The expansions for $\bar{\eta}$ and $\bar{\zeta}$, of which the first two terms are given above, can be shown to be asymptotic, but we cannot take up this question here. The numerical accuracy of (3.632) and (3.633) will be considered in § 4.33.

3.64. The general solution of (3.612) and (3.613) may be put in the form

$$(3.641) \qquad \eta = K_1\eta_1 + K_2\eta_2 + K_3\eta_3 + \bar{\eta},$$

$$(3.642) \qquad \zeta = K_1\zeta_1 + K_2\zeta_2 + K_3\zeta_3 + \bar{\zeta},$$

where K_1, K_2, K_3 are arbitrary complex constants and $\eta_1, \ldots, \zeta_1, \ldots$, have the values determined in the last section.

The particular integral $\bar{\eta}$, $\bar{\zeta}$ represents the motion in an actual trajectory in which ζ is initially zero, and η and η' start with what may be called their equilibrium values, which are numerically very small. The solution $(K_3\eta_3+\bar{\eta})$, $(K_3\zeta_3+\bar{\zeta})$ represents the motion, of the same type, in a trajectory whose initial tangent makes an angle (determined by K_3) with the initial tangent of the chosen plane trajectory. This can be seen from the following considerations. The motion in a slightly different *plane trajectory* would be obtained by omitting all terms in η from the equations of motion of the centre of gravity, and ignoring the equations of angular motion. Equation (3.612) then reduces to $\zeta' = 0$; this represents a trajectory which only differs from a varied *plane trajectory* on account of terms omitted in § 3.21, whose retention renders the equation non-linear. The value of η_3 in (3.627) gives the alteration, through the change in direction of projection, of the first term in $\bar{\eta}$.

---

* *Loc. cit.*, p. 340. We hope to publish these extensions in another place.

3.65. In the usual practical case, the initial conditions take the form

$$\zeta_0 = 0, \qquad \eta_0 = a, \qquad \eta'_0 = b\Omega,$$

where a and b are arbitrary complex constants. It is desirable in such a case to know the degree of importance of the three standard solutions.

The initial values of the standard solutions (retaining the highest order terms only) are as follows:—

$$\begin{array}{lll}
\eta_1 = 1, & \zeta_1 = O(1/\Omega), & \eta'_1 = \tfrac{1}{2}i\Omega(1+\sigma), \\
\eta_2 = 1, & \zeta_2 = O(1/\Omega), & \eta'_2 = \tfrac{1}{2}i\Omega(1-\sigma), \\
\eta_3 = O(1/\Omega), & \zeta_3 = 1, & \eta'_3 = O(1/\Omega), \\
\bar{\eta} = O(1/\Omega), & \bar{\zeta} = 0, & \bar{\eta}' = O(1/\Omega).
\end{array}$$

The constants K_1, K_2, K_3 are determined by the equations

$$K_1\eta_1 + K_2\eta_2 + K_3\eta_3 + \bar{\eta} = a,$$

$$K_1\eta_1' + K_2\eta_2' + K_3\eta_3' + \bar{\eta}' = b\Omega,$$

$$\text{(3.6501)} \qquad K_1\zeta_1 + K_2\zeta_2 + K_3\zeta_3 + \bar{\zeta} = 0.$$

Retaining only the highest order terms these reduce to

$$\text{(3.651)} \qquad K_1 + K_2 = a,$$

$$\text{(3.652)} \qquad \tfrac{1}{2}i(1+\sigma)K_1 + \tfrac{1}{2}i(1-\sigma)\kappa_2 = b,$$

$$\text{(3.653)} \qquad K_3 + O(1/\Omega) = 0.$$

It follows at once that $K_3\eta_3$ is completely negligible compared to $K_1\eta_1$ and $K_2\eta_2$, and that in investigating η we may ignore the third solution (and the particular integral) altogether. On the other hand, the contributions of all the solutions to ζ are of the same order of importance. We shall therefore take as the solution satisfying the most general initial conditions—

$$\text{(3.654)} \qquad \eta = K_1\eta_1 + K_2\eta_2,$$

where K_1 and K_2 are determined by (3.651) and (3.652), and

(3.655)
$$\zeta = K_1\zeta_1 + K_2\zeta_2 + K_3\zeta_3 + \int_0^t \kappa\eta\, dt/c,$$

where by (3.6501),

(3.656)
$$K_3 = -K_1(\zeta_1)_0 - K_2(\zeta_2)_0.$$

§ 3.7. *The Solution of Equations of Type* γ.

The equations of type β are only soluble numerically by step-by-step integration, and will not be considered here, but the equations of type γ (§ 3.4) reduce, when μ is constant and damping effects are neglected, to the equations of a spinning top, and it is convenient to summarize here their solution, in terms of elliptic functions, in the form which is most suitable for our purposes. We shall only consider the initial conditions $\delta = 0$, $\delta' = b\Omega$; this is the rosette form of motion (§ 1.3) and is usually a good approximation to the true motion in its earliest stage. In this case we obtain from (3.404) and (3.405)

(3.701)
$$\phi' = \Omega/(1+\cos\delta),$$

(3.702)
$$\delta'^2 \sin^2\delta - \Omega^2 b^2 \sin^2\delta + \Omega^2(1-\cos\delta)^2 - (\Omega^2/2s)(1-\cos\delta)\sin^2\delta = 0.$$

If we take Ωt as independent variable, the motion depends only on two parameters, b and s. The solution of (3.702) is given by

(3.703)
$$\sin\tfrac{1}{2}\delta = \sin\tfrac{1}{2}\alpha \,\mathrm{cn}\,(K - \lambda\Omega t, k),$$

where α, λ, and k are given by the formulæ

(3.7041)
$$\sqrt{s} = \cos\tfrac{1}{2}\alpha \cosh\tfrac{1}{2}c,$$

(3.7042)
$$b = \tan\tfrac{1}{2}\alpha \tanh\tfrac{1}{2}c,$$

(3.7043)
$$\tan\epsilon = \sin\tfrac{1}{2}\alpha/\sinh\tfrac{1}{2}c, \qquad (k = \sin\epsilon),$$

(3.7044)
$$\lambda = (\sin\tfrac{1}{2}\alpha)/2k\sqrt{s},$$

and K is the complete elliptic integral of the first kind to modulus k. Thus the yaw oscillates between the values 0 and α, and the value of the period T—the interval between successive zeros—is given by

(3.705)
$$\Omega T = 2K/\lambda.$$

The curve of yaw, δ, plotted against Ωt is initially concave (convex) upwards, when $s < 1$ (> 1). This corresponds to the case of instability (stability) for small oscillations.

In the practical analysis of the results of rounds whose stability factor is less than or near 1, it is convenient to use graphical methods. If the observed yaw is plotted against Ωt, it is easy to read off the observed values of α, the maximum yaw, and ΩT, the period. A chart was therefore constructed with suitable families of curves, according to (3.7041)–(3.7044), from which, when ΩT and α are known, s, b, c, and k can be read off directly.

PART IV.—ANALYSIS OF THE EXPERIMENTAL RESULTS.

§ 4.0. *Equations of Motion in Polar Co-ordinates.*

The theoretical results of Part III. will now be applied to the analysis of the observations described in Part II., which consist of determinations of yaw δ and orientation of yaw ϕ, for a shell fired horizontally over a range of about 600 feet. When the stability factor is greater than about 1·1, the maximum yaw for the corresponding round never exceeds 7 degrees, and it is then possible to make use of the complementary function solution of equations of type α as given in § 3.6. These rounds give more valuable information than those which are less stable.

We treat certain of the force coefficients as constants over the range of the experiments, and verify that the results of the theory agree with experiment when certain values are given to the force coefficients. In particular the spin is treated as constant. The way in which the coefficients vary with the velocity is determined mainly by firing shells with various muzzle velocities. The final results have been already described in § 1.2 above.

The experiments determine the values, at definite time intervals along the range (§ 2.0), of the angle of yaw δ and the angle ϕ turned through by the line in which the plane of yaw meets the cards. The measured value of ϕ is zero, when this line is vertical and increases from 0 to 2π radians in the direction in which the shell is spinning. It is, of course, ambiguous by an integral multiple of 2π. Except where specially stated the yaw δ is assumed to be an essentially positive quantity. When OA passes through the position OP, the yaw vanishes; the value of ϕ will change discontinuously by an amount $\pm\pi$, and $d\delta/dt$ will change its sign discontinuously.

It is convenient in Part IV. to express the solution of the equations of motion of type α in terms of the co-ordinates δ and ϕ. The exact relations between the measured δ and ϕ and the direction cosines (l, m, n) and (x, y, z) of § 3.2 are

$$\cos\delta = lx+my+nz,$$

$$\tan\phi = \frac{(nx-lz)\cos\theta_1-(ny-mz)\sin\theta_1}{mx-ly},$$

where θ_1 is the inclination to the horizontal of the tangent to the plane trajectory. Since $\theta_1 < 1\frac{1}{2}$ degrees, we may replace the latter by

$$\tan\phi = (nx-lz)/(mx-ly).$$

Since η is defined by the equation

$$\eta = (m-y)+i(n-z),$$

we obtain, when δ is sufficiently small,

$$\eta = \sin\delta e^{i\phi};$$

this expression neglects terms of the second order compared to those retained. It is an adequate approximation provided $\delta < 7$ degrees.

The general solution for the equations of type α, given in § 3.65, equations (3.654) and (3.6234), is*

$$\eta = (\sigma_0/\sigma)^{\frac{1}{2}}(\mathrm{K}_1 e^{i\mathrm{P}_1}+\mathrm{K}_2 e^{i\mathrm{P}_2}),\dagger \tag{4.01}$$

if we ignore, as we may, the particular integral and the third solution. We shall write

$$\mathrm{P}_1 = p_1+iq_1+p_2+iq_2,$$
$$\mathrm{P}_2 = p_1+iq_1-(p_2+iq_2).$$

Then

$$p_1 = \tfrac{1}{2}\int_0^t \Omega\,dt = \tfrac{1}{2}\Omega t, \qquad p_2 = \tfrac{1}{2}\int_0^t \Omega\sigma\,dt, \tag{4.011}$$

$$q_1 = \tfrac{1}{2}\int_0^t (h+\kappa)\,dt, \qquad q_2 = \tfrac{1}{2}\int_0^t (h-\kappa+2\gamma)\,dt/\sigma, \tag{4.012}$$

and $\sigma^2 = 1-1/s$. We observe that p_1, p_2, q_1, q_2 are all nearly proportional to the time t.

The general solution (equation (4.01)) contains two complex arbitrary constants or four real constants. By a suitable choice of origin for t and ϕ these may be reduced to two. If the time $t = 0$ corresponds to a minimum of δ and the value $\phi = 0$, equation (4.01) may be written

$$\eta = \mathrm{J}\,(\sigma_0/\sigma)^{\frac{1}{2}} e^{ip_1-q_1}\{\cos p_2 \sinh(j-q_2)+i\sin p_2\cosh(j-q_2)\}, \tag{4.02}$$

---

* Treating N and Ω as constant, *i.e.*, neglecting the spin reducing couple Γ.

† Equation (4.01) reduces approximately to the form $\eta = \mathrm{K}t$, when $s = 1$, and to the form $\eta = (\sigma_0/\sigma)^{\frac{1}{2}}\{\mathrm{K}_1 e^{\phi_1}+\mathrm{K}_2 e^{\phi_2}\}$, when $s < 1$, and the shell unstable, the principal parts of ϕ_1, ϕ_2 being real and positive. The solution then fails completely as an approximation to the actual motion except over a small part of the first period. As s approaches the value unity from above, the errors from this cause will begin to increase, but the magnitude of these errors can only be estimated by comparison with the solution of equations of type γ, see § 4.3 below.

where J and j are new arbitrary constants; of these j is small if $|\eta|$ is small at $t = 0$.* The motion is a combination of the following components:—

(1) A uniform rotation about the origin, represented by the term e^{ip_1}.

(2) A damping of the amplitude, represented by $(\sigma_0/\sigma)^{\frac{1}{2}} e^{-q_1}$.

(3) An oscillation of period determined by p_2 whose phase is continually changed by the factor $(j-q_2)$. The values of δ and ϕ are given by the equations

$$\delta^2 = \tfrac{1}{2}\mathrm{J}^2 (\sigma_0/\sigma) e^{-2q_1} \{\cosh 2(j-q_2) - \cos 2p_2\}, \tag{4.031}$$

$$\phi = \phi_0 + p_1 + \operatorname{arc\,tan} \{\coth (j-q_2) \tan p_2\}.\dagger \tag{4.032}$$

So long as $(j-q_2)$ does not change sign, the average rate of increase of ϕ over any number of complete periods is $(p'_1 \pm p'_2)$.

Let α and β be the successive maximum and minimum values of δ (assumed positive). In determining the values of α, β, and the corresponding values of t, it is legitimate to neglect the changes of q_1, q_2, and σ, which are very small in a single period p_2. The maxima and minima are then given by putting $\cos 2p_2$ equal to -1 and $+1$ respectively in (4.031). Writing

$$\alpha_1 = \mathrm{J} (\sigma_0/\sigma)^{\frac{1}{2}} e^{-q_1} \cosh (j-q_2), \tag{4.041}$$

$$\beta_1 = \mathrm{J} (\sigma_0/\sigma)^{\frac{1}{2}} e^{-q_1} \sinh (j-q_2), \tag{4.042}$$

so that α_1, β_1 are defined for all values of t, we have

$$\alpha = \alpha_1 (\mathrm{T}_n), \tag{4.051}$$

for values of T_n given by

$$p_2 (\mathrm{T}_n) = \tfrac{1}{2}(2n+1)\pi, \tag{4.052}$$

$$\beta = |\beta_1 (\mathrm{T}'_n)|, \tag{4.053}$$

for values of T'_n given by

$$p_2 (\mathrm{T}'_n) = n\pi. \tag{4.054}$$

An alternative expression for ϕ is then

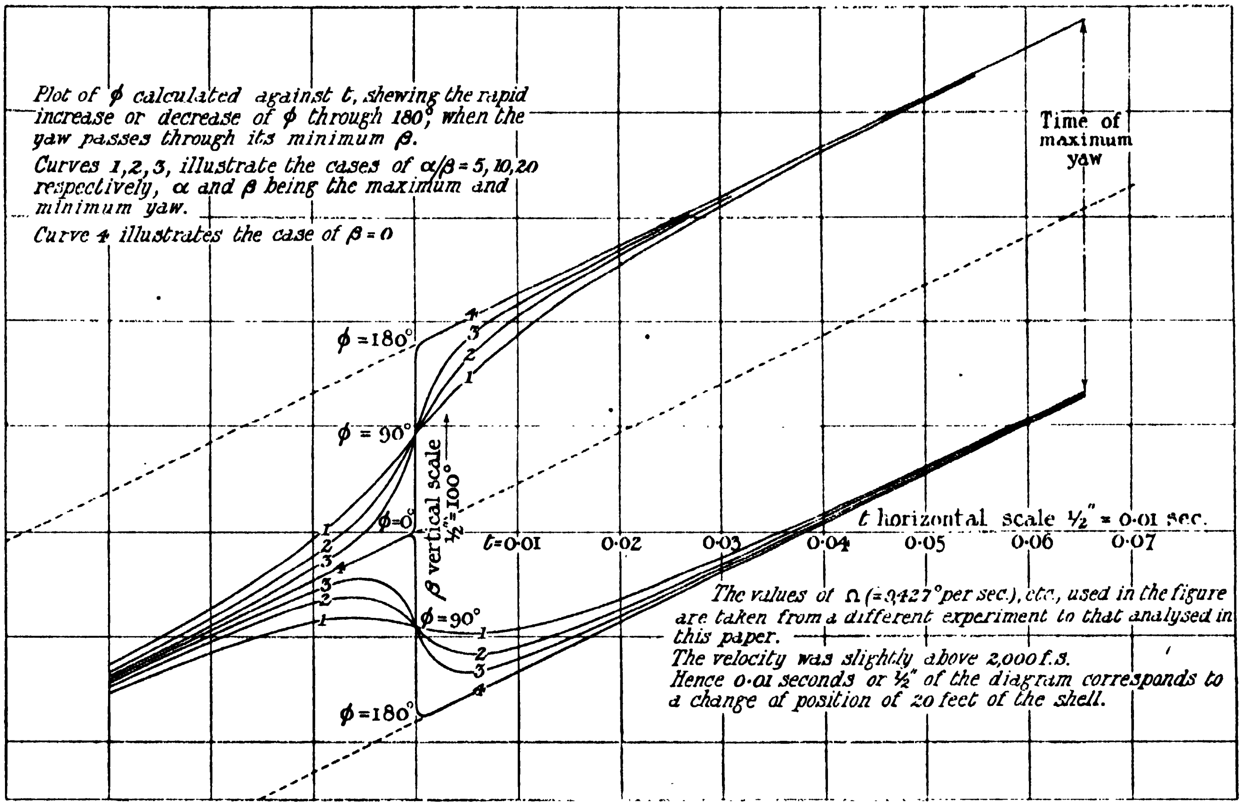
$$\phi = \phi_0 + p_1 + \operatorname{arc\,tan} \left(\frac{\alpha_1}{\beta_1} \tan p_2\right). \tag{4.06}$$

---

* The curves of δ against t appear to have a minimum very near the muzzle of the gun in all rounds fired, but it will be seen that, in analysing the results, it is not necessary to assume any definite origin for t or ϕ.

† Here arc tan (A tan x) is determined in such a way that it changes continuously as x increases indefinitely.

The curves of fig. 11 were calculated from formula (4.06), assuming α_1 and β_1 constant, and p_1, p_2 proportional to t. They show the type of curve on which the observed values of ϕ may be expected to lie.


Plot of ϕ calculated against t, shewing the rapid increase or decrease of ϕ through 180°, when the yaw passes through its minimum β.

Curves 1, 2, 3, illustrate the cases of $\alpha/\beta = 5, 10, 20$ respectively, α and β being the maximum and minimum yaw.

Curve 4 illustrates the case of $\beta = 0$

The values of Ω (= 9427° per sec.), etc., used in the figure are taken from a different experiment to that analysed in this paper.

The velocity was slightly above 2,000 f.s.

Hence 0·01 seconds or ½" of the diagram corresponds to a change of position of 20 feet of the shell.

Fig. 11.

§ 4.1. *Analysis of the Experimental Results.*

It is now necessary to make use of these results to analyse the experiments. The analysis was carried out by graphical methods. The observed values of δ and ϕ were plotted on separate diagrams, examples of which are shown in fig. 12, against the abscissa Ωt, Ω being determined from the muzzle velocity and the observed moments of inertia. The constant factor Ω was inserted to make the independent variable of zero dimensions; the values of the variable Ωt, at given distances down the range, are also independent of small changes in the muzzle velocity. The observed values of δ are sufficient to give a good determination of curves showing the relation between δ and Ωt, except in the neighbourhood of the minima β, where rapid changes of curvature occur when β is small.* These curves give approximate values of the periods from minimum to minimum, and also the best determination available of the values and times of occurrence T_n of the maxima α. By drawing smooth (non-periodic) curves through the values of α we determine α_1 as a function of Ωt.

* When β is small it may be convenient, in a preliminary plot, to change the sign of δ in each alternate period, so as to obtain smooth curves with δ passing through zero periodically.

In drawing curves for ϕ against Ωt it is necessary to resolve the ambiguities of amount $2n\pi$ as follows:—Equation (4.06), or fig. 11, shows that ϕ increases by $\frac{1}{2}\Omega T \pm \pi$ in each period T, the increase of $\pm\pi$ occurring chiefly in the neighbourhood of the minimum, especially when β_1/α_1 is small. Hence, by adding to the observed ϕ

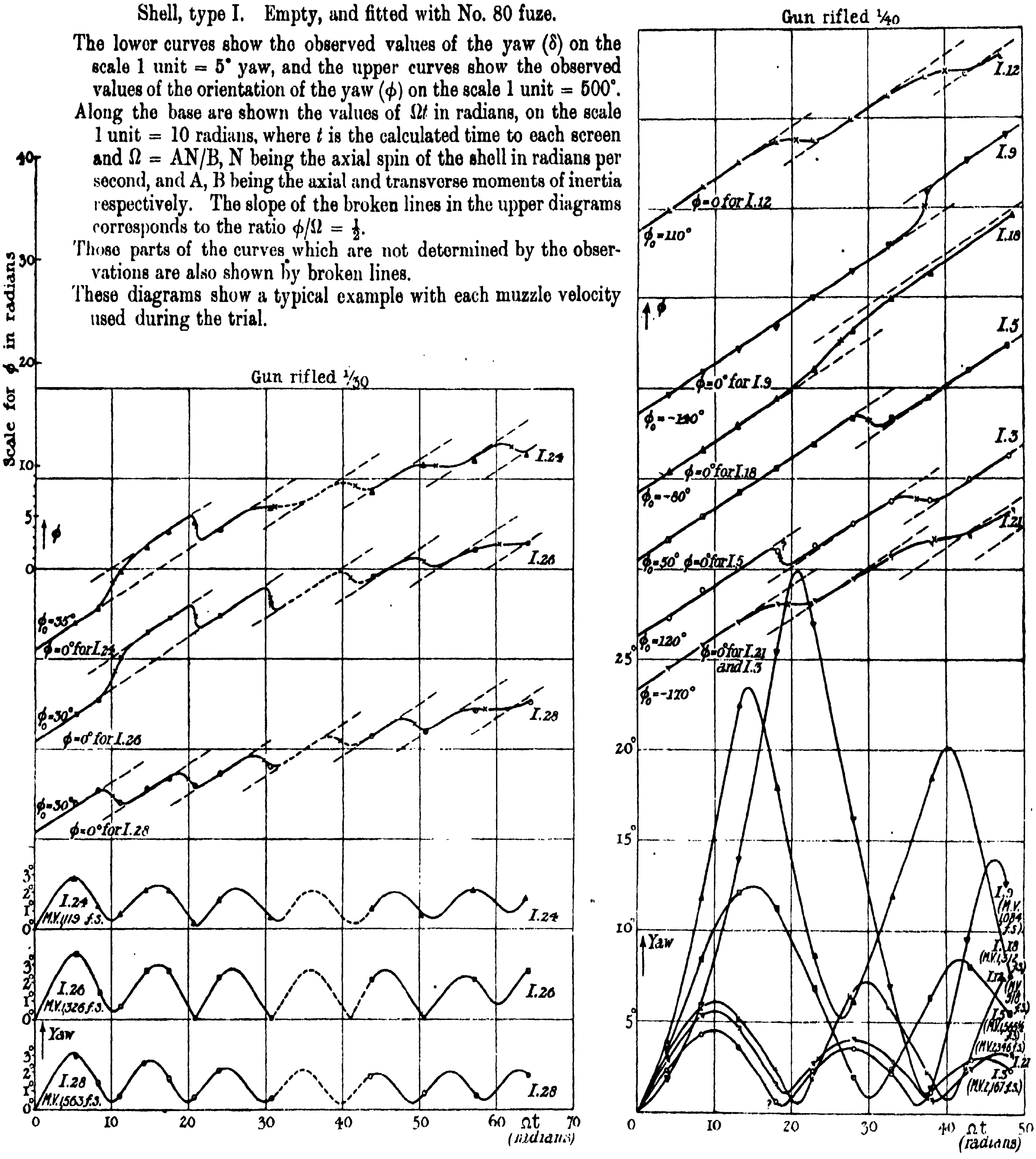
Fig. 12A. Jump card trial, January and February, 1919.
3-in. 20-cwt. gun with 16-lb. H.E. shell, Mark IIB.
Shell, type I. Empty, and fitted with No. 80 fuze.

The lower curves show the observed values of the yaw (δ) on the scale 1 unit = 5° yaw, and the upper curves show the observed values of the orientation of the yaw (ϕ) on the scale 1 unit = 500°.

Along the base are shown the values of Ωt in radians, on the scale 1 unit = 10 radians, where t is the calculated time to each screen and Ω = AN/B, N being the axial spin of the shell in radians per second, and A, B being the axial and transverse moments of inertia respectively. The slope of the broken lines in the upper diagrams corresponds to the ratio $\phi/\Omega = \frac{1}{2}$.

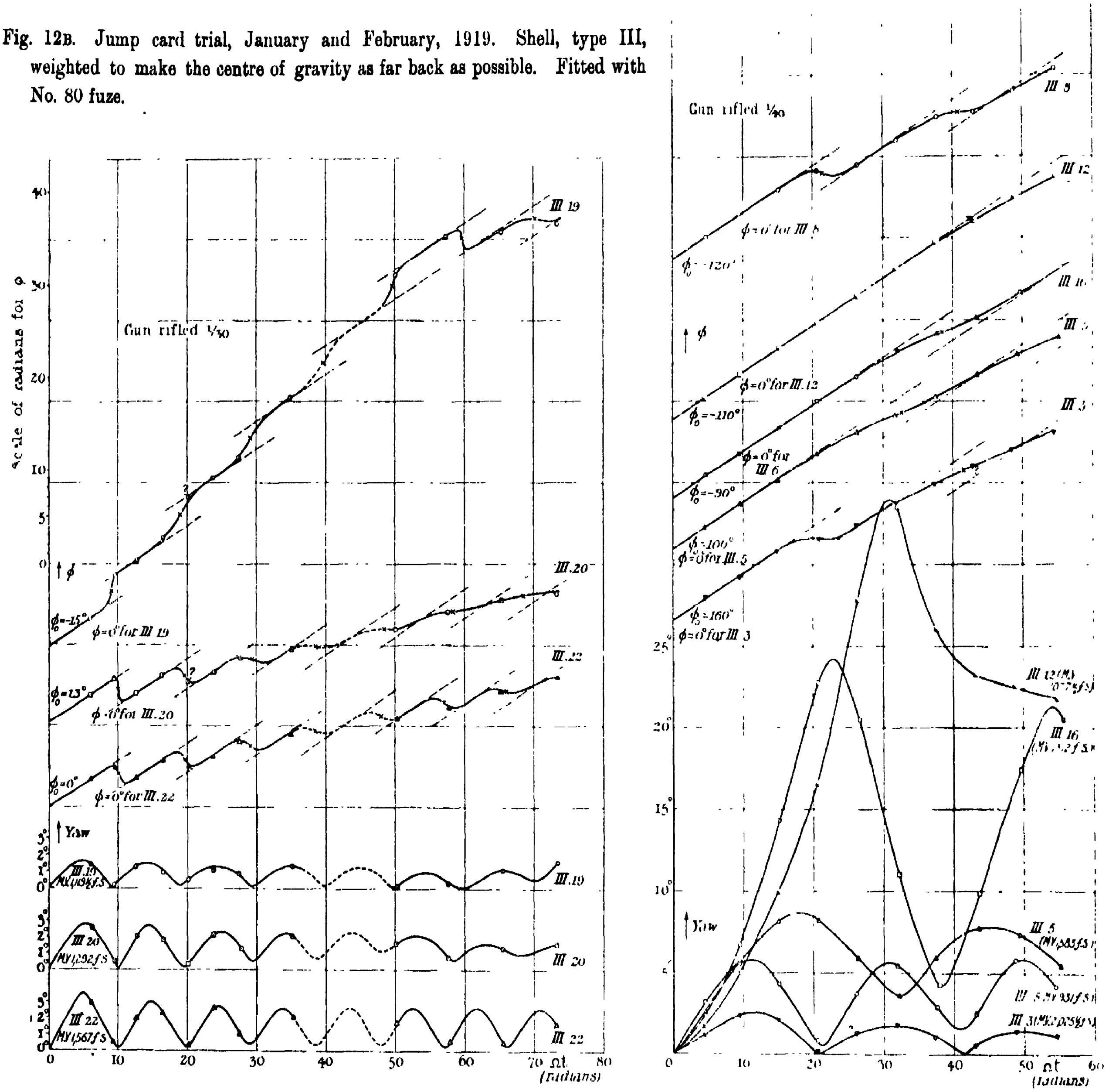
Those parts of the curves which are not determined by the observations are also shown by broken lines.

These diagrams show a typical example with each muzzle velocity used during the trial.



multiples of π, which are alternatively odd and even in successive periods of δ, the points can be fitted roughly to a straight line of constant slope. All the points will lie fairly well on this straight line, except those in the immediate neighbourhood of the minima of δ. By producing this straight line backwards we can determine the initial value of ϕ. The slope of the straight line determines an independent value of Ω, which is equal to the value deduced from the muzzle velocity if the slope is 0·5.

Fig. 12B. Jump card trial, January and February, 1919. Shell, type III, weighted to make the centre of gravity as far back as possible. Fitted with No. 80 fuze.



In practice the values of Ω, obtained by the two methods, were in satisfactory agreement, except for the shells with centre of gravity forward, whose dynamical constants were considerably altered by the set back of the lead block on firing. For these shells the slope of the observed ϕ-curve was taken as defining Ω. The value of B after firing was deduced from this value, and the position of the centre of gravity was determined by equations (2.21) and (2.22).

The curve showing the true relation of ϕ to Ωt must pass through the true values of ϕ, which differ from the observed values only by integral multiples of 2π. It remains doubtful whether the value of ϕ increases or decreases by the amount π radians in passing through a minimum of δ, in addition to its steady increase at rate $\frac{1}{2}\Omega$. This question is settled by the divergence of points near the minimum of δ from the straight line of fig. 11; thus, if the points lie above the straight line in approaching a minimum, there will be an increase of amount π, and *vice versâ*. In this way a continuous curve may be drawn which is consistent with the equation (4.06). Specimens of the curves obtained in the analysis of the actual observations are shown in fig. 12. The portions of the curves in the neighbourhood of the maxima will then coincide approximately with a series of parallel straight lines at distances apart of π radians. The method can only fail in one case when none of the points diverge appreciably from the straight line of fig. 11. This indicates that the value of the minimum β is indistinguishable from zero, while the value of ϕ changes almost discontinuously by $\pm\pi$ at the time of the minimum. It is then immaterial whether the change is taken to be positive or negative.*

The observed values of ϕ in the neighbourhood of the minimum also yield information as to the value of β_1/α_1 and the instant at which the minimum occurs. Let P be any observed point on a ϕ-curve which diverges measurably from the nearest straight portion of the ϕ-curve; lying above it by Δ degrees. Let t_0 be the time of occurrence of the nearest minimum, and δp_2 the change in p_2 between the minimum and P. Then, by (4.06),

$$(4.101) \qquad \cot\Delta = \frac{\alpha_1(t_0)}{\beta_1(t_0)}\cot\delta p_2.$$

If, in this equation, Δ, t_0, δp_2, and $\alpha_1(t_0)$ are regarded as known, we can at once obtain a value of β. By adjusting the value of t_0 we attempt to reconcile the one or more values of β obtained in this manner and also the value demanded by the δ-curve. By combining all the available evidence in this manner, remembering that the δ-curve is nearly symmetrical about a minimum, and the ϕ-curve at the same time halfway between two straight portions, we can draw fairly precise final curves, obtaining values of $\beta_1(t)$ and the times of occurrence of the minima with some accuracy. Such curves are shown in fig. 12.

---

* The rapid changes or discontinuities in the values of ϕ and δ', which occur when δ is very small or zero, are due to the singularity which occurs at the origin of polar co-ordinates. The motion of the shell is, of course, in all cases continuous.

The following quantities have now been determined from the observations, viz.: the (assumed constant) value of Ω, the times T'_1, T'_2, &c., of the occurrence of the minima of δ (those are more accurately determined than the times of the maxima), and the values of $\alpha_1(t)$ and $\beta_1(t)$ over the range of the experiments. These values are given in Table V.

4.11. *Derivation of the Various Force Components.*—It remains to derive the values of the various force components. By equations (4.054), (4.011)

$$(4.111) \qquad p_2(T'_n) - p_2(T'_{n-1}) = \pi, \qquad \int_{T'_{n-1}}^{T'_n} \Omega\sigma\, dt = 2\pi,$$

giving, as a sufficient approximation,

$$(4.112) \qquad \Omega\sigma = 2\pi/T \qquad (T = T'_n - T'_{n-1}),$$

$$(4.113) \qquad s = \frac{1}{1-(2\pi/\Omega T)^2},$$

$$(4.114) \qquad f_M\left(\frac{v}{a}\right) = \frac{A^2N^2}{4Bs\rho v^2 r^3},$$

where σ, s, and v correspond to the time $\frac{1}{2}(T'_n + T'_{n-1})$. T is therefore the time between successive minima of δ. The values of s and f_M obtained in this manner, or, in a similar way, taking an average over several periods, with the corresponding values of μ and v/a, are given in Table VI.,* and provide the data on which figs. 4 and 5 and Table I. were constructed.

By comparing the values of f_M for shells of form A, with three different positions of the centre of gravity, the values of f_L were deduced by the formulæ of § 1.13. This deduction was done graphically as shown in fig. 13. According to § 1.13 the relation between f_M and l, the distance of the centre of gravity from the base of the shell, should be linear. Fig. 13 shows that all the observed points lie on straight lines within the limits of error of the observations. The slope of each line determines the value of f_L The values of f_L are shown plotted against v/a in fig. 4.

---

* For the rounds fired from the gun rifled 1 in 30 the time of the first minimum near the muzzle is, in general, badly determined, and the first period is therefore omitted in determining a mean value for s. For the rounds fired from the gun rifled 1 in 40 the time of the first minimum can be determined with fair accuracy by extrapolation.

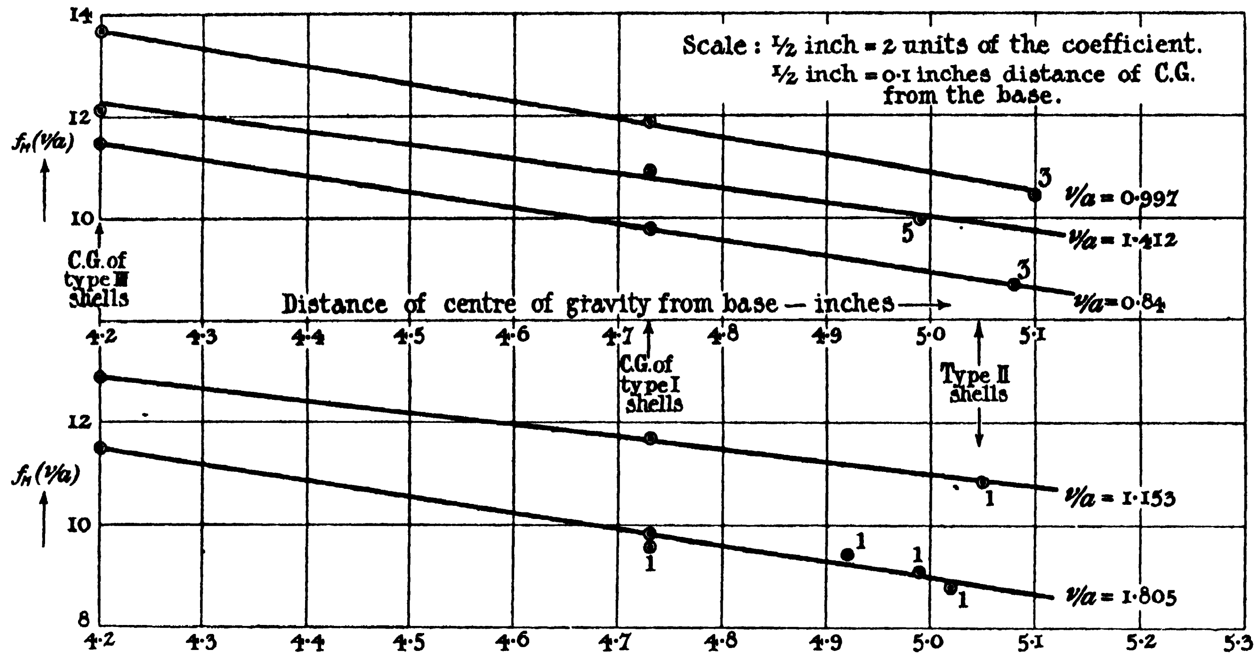


Fig. 13. The determination of the coefficient of the force acting normal to the shell.

The plotted points show the observed values of the couple coefficient plotted against the distance of the centre of gravity from the base.

The slopes of the lines drawn determine the coefficient of the normal force.

The numbers against the points for the Type II. shells give the number of observations whose mean is represented by the plotted point.

4.12. *The Damping Factors.*—It now only remains to derive as much information as possible as to the damping factors κ, h, and γ from the observed values of α_1 and β_1. The factor κ is known in terms of the value of f_L, since, by § 3.1,

$$\kappa = \rho v r^2 f_L/m. \tag{4.121}$$

Squaring and subtracting equations (4.041) and (4.042), we obtain

$$\sigma_0 J^2 e^{-2q_1} = \sigma(\alpha_1^2 - \beta_1^2),$$

$$q_1 = -\tfrac{1}{2}\log\{\sigma(\alpha_1^2-\beta_1^2)\} + \text{const.},$$

$$h+\kappa = \frac{-1}{(t_2-t_1)}\Big[\log\{\sigma(\alpha_1^2-\beta_1^2)\}\Big]_{t_1}^{t_2}. \tag{4.122}$$

In this formula, as well as in those which follow, κ, h, and γ may be treated as sensibly constant over the whole range of one experiment. On dividing (4.041) by (4.042), we obtain

$$\tanh(j-q_2) = \beta_1/\alpha_1. \tag{4.123}$$

Since j and q_2 are both small over the range of the experiments, the formula becomes

$$(4.124) \qquad h-\kappa+2\gamma = \frac{-1}{(t_2-t_1)}\left[\frac{\beta_1}{\alpha_1}\right]_{t_1}^{t_2}.$$

The three equations (4.121), (4.122), (4.124) for κ, h, and 2γ are in theory sufficient to determine their values completely. It may be noted again that 2γ is probably negligible and $h-\kappa+2\gamma$ is always positive, so that q_2 continually increases with the time, *and* β_1/α_1 *continually decreases.* The constant j is always very small, but may be positive or negative. If it is positive, β_1 is initially positive, giving the larger average rate of increase of ϕ, which changes to the smaller rate of increase when β_1 becomes negative. If j is negative, ϕ increases at the slower rate from the beginning. Exactly the opposite results would be obtained if $h-\kappa+2\gamma$ were negative. The values of $h+\kappa$ and $h-\kappa+2\gamma$, obtained in this manner, are given in Table VII.

In order to illustrate the actual path traced out by the axis of the shell, it is necessary to plot δ and ϕ as polar co-ordinates. This is done for three rounds in fig. 14. The resulting curves are roughly equivalent to the path of a point on the axis of the shell relative to the centre of gravity. They illustrate the decrease of α_1, the algebraic decrease of β_1, and the tendency to change from quick to slow precession and to settle down to a *steady* slow precession.

The process described above was evolved gradually during the work of analysing the results, so that a number of observations were analysed before it was fully developed. It is probable that if the calculations were to be repeated *ab initio* a number of periods and minima of δ would be slightly altered, but it is unlikely that any serious systematic errors remain.

4.13. *Details of Tables V. to VII.*—The information contained in the General Table of Results, Table V., has been compiled by analysis of the original standard diagrams. As first constructed these were drawn with the time t as abscissa and not Ωt as in fig. 12. It contains practically all the information of importance provided by the more stable shells. In the unstable cases, a number of which occurred during the trial (see for example fig. 12), a detailed study of the whole yaw curve is required which will not be undertaken in this paper.

Column 5 gives the values of the periods of the yaw curve in units of $\frac{1}{1000}$ second. The periods are read off from positions of the minima and sometimes of the maxima. They are entered to the nearest $\frac{1}{2000}$ second. They are in doubt by more than this quantity in many cases, but mainly in the case of the longer periods, in which small errors are of less importance.

Column 6 gives the values of the maxima of the yaw in degrees and decimals to one place of decimals. These values are read straight from the curves and represent roughly the accuracy to which the maxima are in most cases determined by the observations.

Column 7 gives two entries. The first is the value of ΩT for each round, where T is the mean value of the observed period and Ω corresponds to the *observed* value of the steady rate of increase of ϕ from column 4.

The second entry in column 7 is the velocity of the shell at the middle point of the range of periods whose mean value T is used to determine ΩT. The stability factor determined by ΩT is taken to correspond to this velocity. Finally, in column 8, the values of $\beta_1(t)$ are given with their proper sign as determined incidentally in the determination of their times of occurrence (§ 4.11).

The effect of the cards on the observed value of the period and on s, is ignored in Tables V. and VI. The results obtained here are corrected for this effect, as far as possible, before use in Table I. The information given in Table VI. is deduced directly from Table V. by the equations of § 4.11. In certain cases where the yaw was large it was checked by use of the chart of § 3.7.

The total percentage spread of the values of s (or μ) in the group is in most cases satisfactorily small. The value of 6·7 per cent. for the high velocity group of type I. shells is probably partly due to the fact that the fuzes of shells 1 to 4 were slightly damaged before firing in forcing the shells into the cartridge cases.

At a velocity of 1580 f.s. results were obtained with guns of both twists of rifling. The couple deduced from the results for the gun rifled one turn in 40 calibres is, in the cases of shells of types I. and III., slightly smaller than that deduced from the other gun. This is to be expected as the stability in this case is nearly critical and the maxima are rather large (one maximum is as much as 13 degrees for a type I. shell). The solution of § 3.6 can hardly be expected to apply. The next term in the expression for μ of the form $\mu_1 \sin^3 \delta$ may be expected to be becoming appreciable here; apparently its sign is such that it will tend to diminish the observed value of μ, in agreement with wind channel observations (fig. 2). For the shells of type II. the maxima of the yaw are small in both guns and the results are in agreement.

No perceptible dependence of s on the maximum yaw among the rounds of any one group has been detected in these tables.

The agreement between the results for the two guns at this velocity, and between rounds with different maxima of the yaw, is therefore a satisfactory confirmation of the theory.

The values of $h+\kappa$ and $h-\kappa+2\gamma$, deduced from the observations as explained in § 4.12, are given in Table VII. Of these, the former is more reliable as it does not depend on $\beta_1(t)$ which is difficult to determine. The actual values vary considerably from round to round, and only mean values for each group are shown. The results are therefore very rough, but they indicate qualitatively the nature of the damping, which may also be studied in figs. 12 and 14. For example, in fig. 14C, the motion starts with $\beta_1(t)$ positive, so that the loop encloses the origin, O, or point of zero yaw. But since $h-\kappa+2\gamma>0$, $\beta_1(t)$ diminishes and has become negative by the second minimum, the loop failing to reach O. As $\beta_1(t)$ diminishes further, the loop shrinks to a cusp at the fourth minimum, and the motion soon becomes indistinguishable from a precession at the slower rate. In the meantime, the maximum yaw $\alpha_1(t)$ decreases steadily.

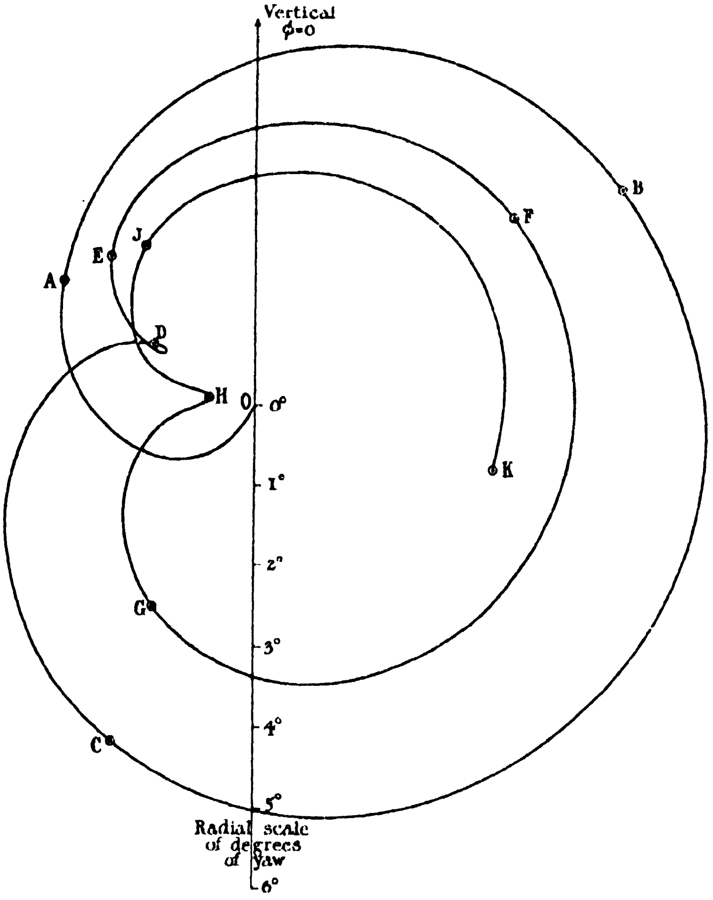


Fig. 14A. Path of nose of shell. Round I.21.

Path described, relative to the centre of gravity, by a point on the axis of the shell in front of the centre of gravity, shown on an enlarged scale.

The total time taken from O to K is 0·2572 second. On the scale used, 1 cm. distance from O represents 1° yaw (very nearly), and corresponds to a linear displacement of 0·118 inch for the nose of the shell from the line of motion of the centre of gravity.

The numerical results for the damping must be affected to some degree by the impacts on the cards, but the available data are not good enough for corrections to be worth making. There is, moreover, the curious phenomenon of an increasing maximum yaw shown by the rounds at 900 f.s. to be accounted for.

The value of κ is known from equation (4.121) and the values of f_L in Table I., so that the damping results determine h and $h+2\gamma$ or, more accurately, $h+2\gamma-\Gamma$ (§ 3.62). It at once appears that $2\gamma-\Gamma$ is negative and of much the same order as h. This is somewhat unexpected. Of course Γ (or $-N'/N$) is positive, but it is hardly likely that its numerical value is much larger than 0·03. It is natural to expect

γ to be small and positive,* which does not fit in with the observations. Further experiments would be needed to throw light on all these points.

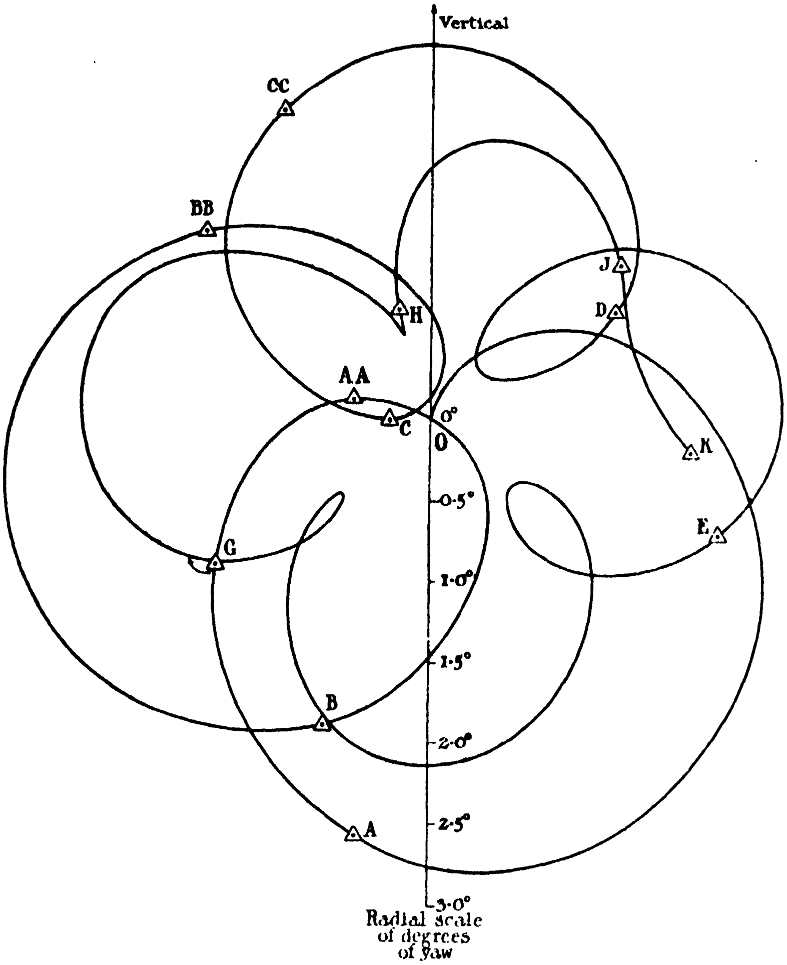


Fig. 14B. Path of nose of shell. Round III.20.

Path described, relative to the centre of gravity, by a point on the axis of the shell in front of the centre of gravity, shown on an enlarged scale.

The total time taken from O to K is 0·3647 second. On the scale used, 2 cm. distance from O represents 1° yaw (very nearly), and corresponds to a linear displacement of 0·128 inch for the nose of the shell from the line of motion of the centre of gravity.

In the fourth column what appears to be the most probable value of h is given; the values of f_H in Table II. are based on these figures and obtained by the equation (see §§ 3.5 and 1.12)

$$f_H = \frac{hB}{\rho v r^4}.$$

---

* The coefficient γ comes from the swerving couple J (§ 1.12). This couple will only arise if the swerving force K does not act through the centre of gravity. Since the air pressures are greater near the nose than near the base, we may expect K to act *in front* of the centre of gravity. By analogy with the connection between the direction of rotation and the direction of the resulting swerve on a golf or tennis ball at low velocities, we may expect K to act along the axis of M reversed in fig. 9, for a right-handed twist of rifling. This would result in a positive value for γ.

The figures in Table VII. were obtained from the sufficiently stable rounds fired from either gun. In the one comparative pair of groups available the results for the two different stability factors and values of Ω were in agreement.

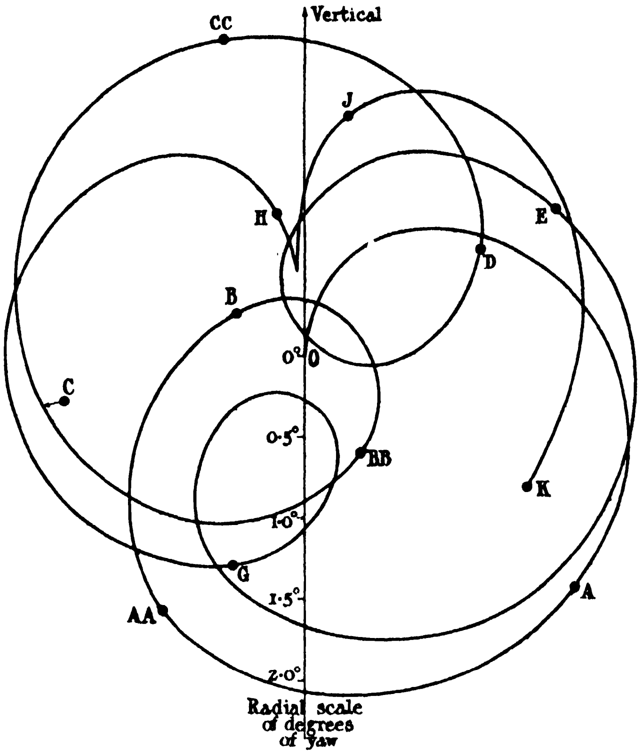


Fig. 14c. Path of nose of shell. Round IV.15.

Path described, relative to the centre of gravity, by a point on the axis of the shell in front of the centre of gravity, shown on an enlarged scale.

The total time taken from O to K is 0·5502 second. On the scale used, 2 cm. distance from O represents 1° yaw (very nearly), and corresponds to a linear displacement of 0·143 inch for the nose of the shell from the line of motion of the centre of gravity.

Note that the first loop encloses O, corresponding to the "stepped up" motion in ϕ. Subsequent loops do not, as the motion in ϕ has changed to the "stepped down" motion. (See fig. 12.)

§ 4.2. *Determination of the Motion of the Shell in Space.*

We now proceed to make use of the results of the experiments to determine the true path of the centre of gravity of a shell projected in a given manner. The solution of the equations of type α is sufficient for this purpose so long as the yaw does not exceed 0·1 radian; the values of f_L, f_M, f_H, &c., which we have obtained, are sufficient to determine the motion completely in this case. Assuming that the maximum yaw due to the initial disturbances is less than 0·1 radian in the first period, it will remain so throughout the trajectory; the yaw arising from the particular integral will not exceed 0·1 radian until the velocity has fallen considerably below 700 feet per second. Hence, when the yaw exceeds 0·1 radian, the wind channel values for the various force components as shown in fig. 2 can be used; it will, however, be necessary to abandon the above method of solution and proceed by the step-by-step integration of the equations of type β.

Throughout the following work all numerical results will be based on a set of plane trajectories of a 16-lb. shell, of external form A, fired at a muzzle velocity of 2000 f.s., calculated by the ordinary ballistic methods.* The various elements of the trajectories at elevations of 30 degrees and 50 degrees and a list of constants for the service shell, to which the calculations apply, are given in Table VIIIA.

From the value of ζ for the general solution, as given in § 3.65, we can deduce the true path of the centre of gravity in terms of the tabulated elements of the plane trajectory. Let $(X_1, Y_1, 0)$ be the co-ordinates of the shell in the plane trajectory at time t, and (X, Y, Z) the corresponding co-ordinates in the true (twisted) trajectory. The direction cosines of the tangents to the two trajectories are $(X'_1/v_1, Y'_1/v_1, 0)$, or $(\cos\theta_1, \sin\theta_1, 0)$ and X'/v, Y'/v, Z'/v, so that, to the usual order of approximation,

$$\begin{aligned} c\zeta &= (Y'-Y'_1)(\cos\theta_1)/v_1 - (X'-X'_1)(\sin\theta_1)/v_1 + iZ'/v_1 \\ &= (H'+iZ')/v_1, \end{aligned} \tag{4.201}$$

say, while the condition $v = v_1$ gives

$$(X'-X'_1)\cos\theta_1 + (Y'-Y'_1)\sin\theta_1 = 0. \tag{4.202}$$

It is convenient to separate the parts of the solution arising from the complementary function and the particular integral. To determine the latter, we use equations (3.632), (3.633), and (4.201), obtaining

$$\begin{aligned} \frac{Z'}{v_1} &= c\int_0^t \frac{-4s\kappa\theta_1'\,dt}{c\Omega} \\ &= c\psi, \end{aligned} \tag{4.203}$$

say, neglecting the terms $i\theta''_1/\Omega$ in Φ (see § 3.20). This equation defines ψ. Therefore

$$Z = \int_0^t \psi v_1 \cos\theta_1\,dt, \tag{4.204}$$

where ψ may be written (since $-\theta'_1/c = g/v_1$)

$$\psi = \int_0^t \frac{g\kappa AN}{\mu v_1}\,dt = \frac{Ag}{mr}\int_0^t \frac{N f_L(v/a)}{f_M(v/a)}\frac{dt}{v^2}.$$

To the same approximation $(X'-X'_1)/v_1$ and $(Y'-Y'_1)/v_1$ are $O(1/\Omega^2)$, so that (X_1-X) and (Y_1-Y) are small compared to Z, so long as the approximations hold. The above result is identical in form with the "classical" formula of MAYEVSKI, freed from the unnecessary restriction that f_L/f_M and N should be constants.* We have thus justified the use of the plane trajectory as an approximation to the true motion. The leading terms in $(X-X_1)$ and $(Y-Y_1)$ can be calculated if required.

---

* Trajectories were calculated with the ballistic coefficient 1·75.

The effect on the motion of a change in initial conditions is obtained from the complementary function. Equation (3.655) gives the value of ζ corresponding to the general initial conditions $\zeta_0 = 0$, $\eta_0 = a$, $\eta'_0 = b\Omega$, where a and b may be complex. Substituting in equation (4.201) the part of ζ arising from the complementary function, it appears that $H+iZ$ is made up of three parts :—

(*a*) A periodic term

$$H_1+iZ_1 = -\frac{4\kappa v_1}{\Omega^2}\left(\frac{K_1\eta_1}{(1+\sigma)^2}+\frac{K_2\eta_2}{(1-\sigma)^2}\right).$$

(*b*) A term

$$H_2+iZ_2 = -\{K_1(\xi_1)_0+K_2(\xi_2)_0\}\int_0^t cv_1\xi_3\,dt,$$

which is the effect of a variation in the direction of projection, as mentioned in § 3.64.

(*c*) A constant term H_3+iZ_3 equal to the initial value of H_1+iZ_1 with its sign changed.

4.21. *Numerical Results as to Motion of Centre of Gravity.*—The only data as to the forces on the shell required for the calculation of the drift are the value of f_L/f_M as a function of v/a. This is derived from the results of the jump card experiments for $v/a > 0{\cdot}7$, and from wind channel experiments for $v/a < 0{\cdot}7$, and is shown plotted in fig. 15.

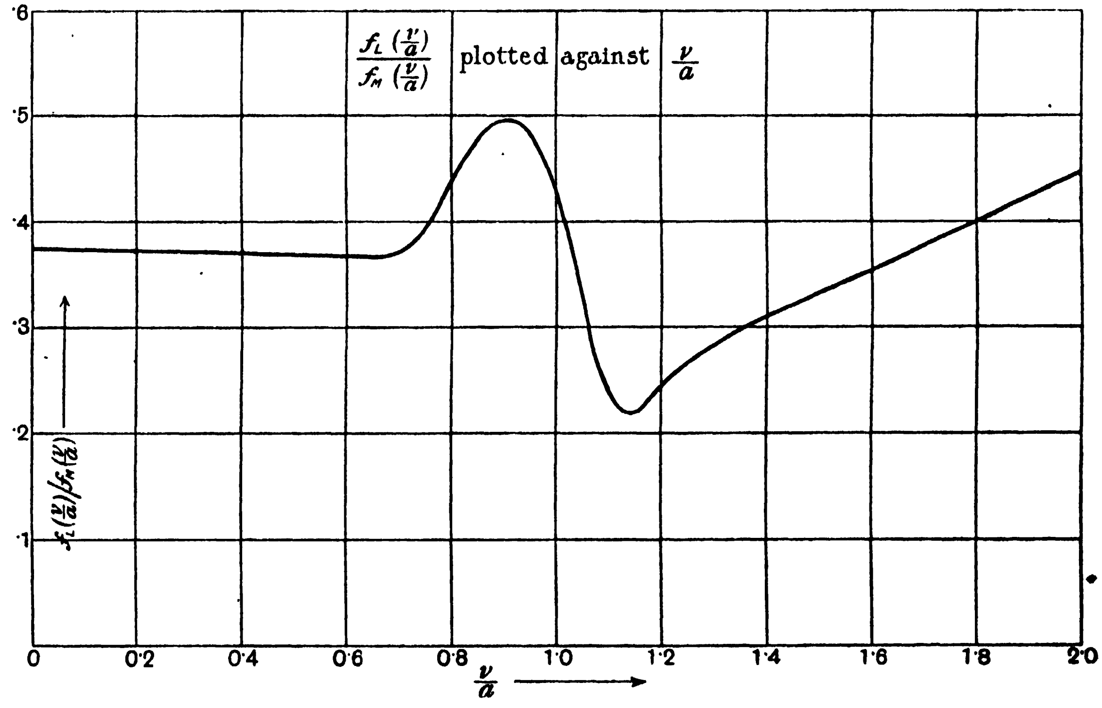


Fig. 15.

---

* Prescott obtains a solution of the equations of motion in the form of a series of which the first term is also equivalent to Mayevski's formula. (See Introduction, p. 296.)

As the value of the couple Γ is only known to be small, it is necessary to assume that N is constant. The principal steps in the calculation of the drift Z, by means of (4.203) and (4.204), for the trajectories at 30 degrees and 50 degrees, are given in Table VIIIb. for the gun rifled 1 turn in 30 diameters of the bore. For a different rifling the drift (N constant) is proportional to N.

It is only necessary to estimate roughly the effect of the complementary function on the motion of the shell since the total effect is always fairly small. The periodic term H_1+iZ_1 is obviously smaller than $4\kappa v_1\alpha_1/\Omega^2(1-\sigma)^2$ in absolute value, where α_1 is defined as in § 4.0, equation (4.041). The initial value of the coefficient of α_1 is 1·25 feet for the gun rifled 1 turn in 30 diameters of the bore; both α_1 and its coefficient diminish rapidly. Taking $\alpha_1 = 0\cdot 1$ radian as an extreme case,

$$|H_1+iZ_1| < 1\cdot 5 \text{ inches.}$$

The actual value in practice is probably always $< 0\cdot 5$ inch, which is small. It explains why no evidence of helical motion was obtained in the jump card experiments. The constant value of $|H_3+iZ_3|$ is equal to the initial value of $|H_1+iZ_1|$ and is also negligible. There remains only the term H_2+iZ_2. This is equivalent to an angular deviation of $|K_1(c\xi_1)_0+K_2(c\xi_2)_0|$, which is less than $2\kappa\alpha_1/\Omega(1-\sigma)$. The coefficient of α_1 for the gun rifled 1 turn in 30 diameters of the bore is $1\cdot 8\times 10^{-2}$, so that for a value of α_1 of 0·1 radian the angular deviation is of the order 0° 6′. This is of the same order of magnitude as the angular jump likely to be due to changes of form and position in the gun and mounting under firing stresses. When it varies from round to round in magnitude and direction, it will account for an *irregularity* of the corresponding amount in the observed positions of the shells at any time. When, as may sometimes be the case, it remains fairly constant from round to round, it will cause systematic errors in the position of the shell at any time. It is probable that anomalous values of the drift, sometimes observed at short times, are due to this cause. Practical results, however, more often fully justify the use of the particular integral alone to give a mean value of the drift when the initial disturbance is only known to be small.

The results of the above calculations of drift will now be compared with observations of the Z co-ordinates of the bursts of shells, fired at Portsmouth, at corresponding elevations, in February and April, 1918. For this purpose use is made of the azimuth of the shell burst Z/X; the quantity $A = Z/Xt$ is tabulated, since its value varies slowly along the trajectory (Table IX.). The agreement between observation and calculation is as good as could be expected, in view of the uncertainty in the wind effects, and provides important evidence as to the correctness of the whole theory.

4.22. *The Damping of the Angular Oscillations and the Effect on the Head Resistance.*—We have now obtained the complete motion of the centre of gravity of the shell by use of the equations of type α for the two standard trajectories; we have, in so doing, assumed that the velocity of the shell is the same in the plane and true trajectories; we must now examine more closely the possible effect on the drag of the angular oscillations and their rate of damping, by means of the values of h and κ obtained above. From equations (4.031) and (4.032) it appears that for sufficiently large values of t, δ and ϕ are given approximately by the equations

$$\delta = \tfrac{1}{2}\mathrm{J}\,(\sigma_0/\sigma)^{\frac{1}{2}}\,e^{-(q_1-q_2)},$$
$$\phi = \phi_0+(p_1-p_2),$$

so that the shell settles down to a steady precession with the slower precessional angular velocity, the yaw gradually diminishing in proportion to the factor $(\sigma_0/\sigma)^{\frac{1}{2}}\,e^{-(q_1-q_2)}$. This quantity is tabulated in column 9, Table VIIIB., on the assumption that $\gamma = 0$ and $h = 3\kappa$. The damping is actually more rapid than is indicated by this approximation.

The question of the rate of damping of the initial oscillations of a shell is of importance on account of its effect on the drag R, for the effect, though it may be small, will be cumulative, since it tends always to increase R. If it is assumed that the effect on R is given by*

(4.221)
$$\mathrm{R} = \mathrm{R}_0\,(1+k\delta^2),$$

where R_0 is a function of v and k is a constant, it is possible to obtain an approximate formula for the total change in velocity produced on the assumption that the time taken to damp out the oscillations is relatively small. We have

(4.222)
$$-\Delta v = \frac{1}{m}\int_0^\infty (\mathrm{R}-\mathrm{R}_0)\,dt,$$
$$= \frac{k\mathrm{R}_0}{m}\int_0^\infty \delta^2\,dt.$$

Using (4.031), and integrating on the assumption that σ is constant, q_1, q_2, and p proportional to t, and j zero, we obtain

(4.223)
$$-\Delta v = \frac{k\mathrm{J}^2\mathrm{R}_0}{2m}\int_0^\infty e^{-2q_1}(\cosh 2q_2-\cos 2p_2)\,dt,$$
$$= \frac{k\mathrm{J}^2\mathrm{R}_0\sigma_0{}^2\,(h_0+\kappa_0)}{2m\,\{\sigma_0{}^2\,(h_0+\kappa_0)^2-(h_0-\kappa_0+2\gamma_0)^2\}}\,.$$

At present we have no information as to the value of k except at low velocities, while J varies from round to round so that no numerical results can be given. It seems likely that this is a cause of irregularities in range in practice of first class importance. The yaw arising from the particular integral will also tend to increase the resistance, but the effect is of less importance in a low angle trajectory.

---

* By symmetry, there can be no odd powers of δ in R.

4.23. *The Exact Motion in a High Angle Trajectory.*—It will be shown in the next section that, for a trajectory of much higher elevation than 50 degrees, the approximations for the particular integral break down, and the equations of type α, are not applicable to the later stages of the trajectory when the velocity has fallen much below 500 f.s. These later stages occur after the initial oscillations have been damped out, and are suitable for the use of equations of type β. These equations can be integrated step-by-step on the basis of the wind channel values of R, L, and M (fig. 2), which apply to velocities up to 700 f.s. The process is analogous to the usual method of calculating a plane trajectory, but rather more laborious, and has been carried out in one case only for a 3-inch $12\frac{1}{2}$-lb. shell fired at 70 degrees with a muzzle velocity of 2450 f.s. At 40 seconds the yaw has reached the large value of 60 degrees and is still increasing. This is partly due to the large initial value of the stability factor (about 4·0) indicating that the spin is unnecessarily large for this shell. The results of comparing the drift with observation were again fairly satisfactory in this case; but details of these results cannot be given here.

§ 4.3. *Estimate of the Errors in the Various Solutions.*

In the development of the various solutions of the equations of motion in Part III., it was found necessary to neglect certain terms. We shall now proceed to examine these terms in succession, and to determine, as far as possible, their numerical values, using the values of the various force components obtained from our experiments. By so doing we shall justify the use of the solutions by showing that the terms neglected are all very small over the range covered by the jump card experiments. In the applications to the later parts of a trajectory, the solutions break down in certain cases, and an examination of the error terms enables us to define the circumstances under which the solutions are valid. We proceed to examine the various terms. It is necessary as a rule to distinguish the terms neglected in obtaining the complementary function from the terms neglected in obtaining the particular integral.

In the complementary function, m, n, y, z are periodic functions of the time with periods comparable with Ω. For the solution to be applicable we have to assume that δ is always small (say $\delta < 0 \cdot 1$ radian). Then m, n, y, z are all small quantities comparable with δ, and m'/Ω, m''/Ω^2, &c., are also comparable with δ, while $(1-l)$, l'/Ω, l''/Ω^2, &c., are of the order of δ^2. In neglecting terms independent of θ'_1 from the equations (3.202), (3.203), we are guided by the condition that all terms neglected should be of the order of δ^2 compared with those retained. As regards the terms containing θ'_1 or θ''_1, it appears that the maximum value of θ'_1/Ω in the 50 degrees trajectory (for rifling 1 in 30) is 30×10^{-5}, its initial value being 5×10^{-5}. Hence all terms such as $nm'\theta'_1$, $n\theta'^2_1$ are completely negligible in obtaining the complementary function.

If all terms in θ'_1, θ''_1 are removed from equations (3.202), (3.203), they become equivalent to the equations of type γ; the errors in neglecting further terms may, therefore, be determined by comparing the solutions of equations (3.204), (3.205), &c., of type α with those of equations (3.404), (3.405), of type γ (assuming μ constant in both cases). Equations (3.7041), (3.7043), (3.7044), (3.705) can be made to give the following approximation to the true value of s in terms of T and α:—

$$s = \frac{1}{1-4\pi^2/\Omega^2\mathrm{T}^2}\{1-\tfrac{1}{8}(2s+1)\alpha^2\}.$$

This is valid so long as (1) α is so small that α^4 may be neglected, and (2) $s-1$ is positive and large compared with α^2. Comparing this with the corresponding first approximation (4.113), we obtain the error in the value of s due to the neglect of the terms in $(1-l)$, l'', &c., in equations (3.202), (3.203). The relative value of the error is given in the following table:—

α	$s = 1 \cdot 1.$	$s = 2.$	$s = 3.$
10°	0·012	0·019	0·027
5°	0·0030	0·0048	0·0066
2·5°	0·0007	0·0012	0·0016

In analysing the jump card trial, whenever the error from this cause is appreciable, the results have been corrected by determining the values of s from the chart described in § 3.7.

It appears also from the solution of the equations of type γ that when $s \leqq 1$ the initial angular motion is still periodic, but no longer of the nature of a small oscillation, since the period is a function of the amplitude and tends to infinity as the initial disturbance tends to zero.

In using equations (3.202), (3.203) to obtain the particular integral, the order of magnitude of the various terms is different. The term $\mathrm{AN}l\theta'_1$ is now the most important, while n is $\mathrm{O}(1/\Omega)$ and m is $\mathrm{O}(1/\Omega^2)$ with the notation of § 3.6. Most of the terms neglected are then $\mathrm{O}(1/\Omega^4)$ compared to the principal term, and completely insignificant, but $\mathrm{B}nl\theta'^2_1$ is $\mathrm{O}(1/\Omega^2)$ and would affect the third term in the expansion for $\bar{\eta}$. Its effect however is completely negligible.

4.31. *The Equations of Motion of the Centre of Gravity.*—These equations may be treated in a similar manner. In obtaining the complementary function, y and z are small compared to m and n (see equations (3.624), (3.625)), κ/Ω being initially less than 0·01. As regards the differential equation for u (3.2141), the effect of neglecting the terms arising from the variation of R with δ has been discussed in § 4.22; no numerical data are available; the effect is theoretically second order. The term in $1-x$ is obviously negligible. Omitting these terms, the equation can be reduced to the form

$$u'+\alpha u = -gy\cos\theta_1,$$

where

$$\alpha = \frac{1}{m^*}\frac{\partial}{\partial v_1}\{\mathrm{R}(v_1+\theta u,\ \theta\delta)\} \qquad (0<\theta<1).$$

As a rough approximation we may assume that $\mathrm{R} = kv^2$, so that $\alpha = -2v'_1/v_1$. We then find that

$$\frac{d}{dt}\frac{u}{v_1^2} = -\frac{yg\cos\theta}{v_1^2}.$$

In the case of the complementary function, y consists of a constant term less than 2×10^{-3}, and periodic terms whose period is of order $1/\Omega$.

The former makes a contribution to u/v_1 which is still less than 10^{-3} after 20 seconds. The latter contributes† a term of order $yg/v_1\Omega$ which is always less than 3×10^{-6}.

In the case of the particular integral y is $\mathrm{O}(1/\Omega^2)$, see § 4.2. Hence in all cases we are justified in putting $u=0$, $v=v_1$, so long as the equations of type α hold at all, with the proviso that this conclusion may be at fault if the k of § 4.22 is numerically large.

In reducing equations (3.212) and (3.213) we put $x=1$, $\cos\delta=1$. This amounts to neglecting $1-x$, $1-\cos\delta$ compared to 1, and is obviously justifiable. We omit altogether from (3.212) the terms $x\theta'_1+(g/v)\cos\theta_1$, or $-g\cos\theta_1\,(x/v_1-1/v)$. This term is excessively small, but could be retained, if desired. Finally we omit the terms in $y\cos\theta_1$, justifying the omission by the arguments used above for the same term in the equation for u.

§ 4.32. *Errors in the Solution for the Complementary Function.*—The second term R_1 in the expansion of R in equation (3.6233) will be taken as representing the principal part of the error in the standard solution for the complementary function arising at this stage. Its value is

$$\mathrm{R}_1 = -\tfrac{1}{4}\mathrm{M}^{-\frac{3}{4}}\frac{d^2}{dt^2}(\mathrm{M}^{-\frac{1}{4}}),$$

where M has the value appropriate to (3.6231). For simplicity in estimating errors we may take only the leading term in M so that here

$$\mathrm{M} = \tfrac{1}{4}\Omega^2\sigma^2.$$

The values of s determined from the jump card trial and the data of the 50 degrees plane trajectory are tabulated in column 2 of Table VIIIB. The value of σ can be deduced and its first and second differential coefficients obtained from a difference formula. In this way we find that initially

---

† This contribution is of the form $\int_0^t f(t)\,e^{i\Omega t}\,dt$, which is of the order $(1/\Omega)\times$ (maximum of $f(t)$) under suitable restrictions on $f(t)$, which are satisfied here.

$$R_1/R_0 = 0{\cdot}000011,$$
$$R_1/R_0 = 0{\cdot}0031,$$

for the guns rifled 1 turn in 30 diameters and 40 diameters respectively. Moreover, the value of R_1/R_0 diminishes along the trajectory. The neglect of this term is therefore justified, provided $s > 1{\cdot}1$, and the total error in the solution will probably be of the same numerical order.

The contribution of ζ (see 3.6235) to the coefficient of η in equation (3.613) is

$$-\frac{2\kappa c'}{c(1 \pm \sigma)},$$

the principal term in this coefficient being $-\Omega^2/4s$. The relative value of the error in omitting this term is therefore

$$\left|\frac{8sc'\kappa}{c\Omega^2(1 \pm \sigma)}\right|,$$

which is less than

$$\left|\frac{4s\kappa}{\Omega}\frac{4s\theta'_1}{\Omega}\tan\theta_1\right|.$$

The values of these three factors can be obtained from Table VIII. The maximum value of this ratio for the 50 degrees trajectory is $\frac{1}{250}\tan\theta_1$. This is negligible.

In evaluating M to obtain equation (3.6234), terms such as κ^2/Ω^2, $\kappa h/\Omega^2$, κ'/Ω^2 are neglected. It is unnecessary to evaluate such terms in detail, since it is known that κ/Ω and h/Ω are less than $0{\cdot}02$ in all cases. It would, however, be easy to write down equation (3.6234) with such terms included.

4.33. *Errors in the Particular Integral.*—The errors of the expression for the motion of the centre of gravity of the shell, given in (3.632) and (3.633), may be obtained from the expansion of the particular integral in powers of $1/\Omega$. The ratio $i\theta''_1/\Omega\theta'_1$ of the two terms in Φ, § 3.2, can be worked out from the data of the plane trajectory. Its initial and greatest value for the gun rifled 1 turn in 30 diameters is $(0{\cdot}0008)$, so that the second term is entirely negligible in comparison with the first. Writing therefore $\Phi_1 = \theta'_1$, it appears that the terms of order $1/\Omega^2$ in (3.632) are real and so do not affect the drift. The next term is (with $\gamma = 0$)

$$\frac{4s}{(i\Omega)^3}\left\{-\eta''_1 + \frac{d}{dt}(4s\eta'_1) + \eta'_1(8s\kappa - h - \kappa) + \eta_1\left[\frac{d}{dt}(4s\kappa) + 4s\kappa^2 - h\kappa - \kappa'\right]\right\},$$

where η_1 $(= 4s\theta'_1)$ is the coefficient of $1/i\Omega$ in the first term in the expansion of $\bar{\eta}$.

There are also a number of other terms involving c', c'' and c'^2. The terms in c' are very small initially and vanish at the vertex, so that they are never likely to become important. The other terms in c'' are certainly very small provided that s is of order unity. Since s varies roughly inversely as the square of the velocity (*i.e.*, f_M constant), the magnitude of the terms containing s rises very rapidly in the later stages of the trajectory when v becomes small. The first term in $\bar{\eta}$, $-4is\theta'_1/\Omega$, is given numerically in Table VIII., where it appears how rapidly it increases as the velocity falls. The values of the second term as given in equation (3.632) are also given (Table VIIIB., column 8). It appears that the ratio of the second term to the first term is always small so long as the first term is small. This term represents the effect of the particular integral in altering the co-ordinates in the plane of fire. The third term as given above is more difficult to evaluate, and only a rough estimate has been made of its value at two points on the 50 degrees trajectory. The results are:—

Seconds.	Third term/first term.	Third term.
$t = 0$	$-2 \cdot 02 \times 10^{-8}$	$-8 \cdot 5 \times 10^{-7}$
$t = 20$	$-1 \cdot 94 \times 10^{-2}$	$-7 \cdot 2 \times 10^{-4}$

The value of the drift as estimated by the first term is therefore slightly too large. The first part of the third term, $-4s\eta''_1/(i\Omega)^3$, is of special interest, as it represents the sole contribution of the term η'' in equation (3.613) to the value of $\bar{\eta}$ to this order. The term η'' represents all that remains in the equations of type α of the terms in B neglected in § 3.3 in obtaining the equations of type β. The initial value of $-4s\eta''_1/(i\Omega)^3$ is only $3{\cdot}46 \times 10^{-5}$ of the first term in $\bar{\eta}$ in the 50 degrees trajectory, and this ratio does not tend to increase as the velocity diminishes. This makes it likely that the equations of type β give an accurate solution in all cases when the initial conditions are those of the particular integral.

Returning to the particular integral, we have shown that the third term is only 0·03 (?) of the first term at the vertex of a 50 degrees trajectory where the velocity is as low as 500 f.s. For a trajectory at still higher elevation the minimum velocity is lower; the value of the first term soon becomes too great for the use of approximations which neglect $1-\cos\delta$, while the third term can no longer be neglected in comparison with the first term. The solution therefore fails when the elevation much exceeds 50 degrees as soon as the velocity has fallen much below 500 f.s. The true degree of approximation given by the expansion can only be obtained in a special case. If the terms in η'' in equation (3.613), and terms of the solution containing c', &c., arising from the terms in ζ are neglected, it may be shown that the error of the expansion at any stage is less in numerical value than the last term retained.[*] Hence the numerical estimates of the third term, obtained above, justify the use of the first term only to obtain an approximate value of the drift at all elevations up to 50 degrees and for the initial part of a trajectory at any elevation.

PART V.—SUMMARY AND CONCLUSION.

§ 5.0. *Summary of Preceding Results.*

In the earlier parts of this paper we have suggested a tentative set of components for the complete force system acting on a shell moving through air (or other medium), in which this complete system may be assumed to depend at any moment only on the position and velocities of the shell. We have submitted these suggestions to the test of experiment, and found that, so far as we have carried the analysis in this paper, the experiments confirm our suggestions, and provide, when the yaw is small, numerical values for two of the force coefficients (f_M with a probable error of 2 per cent. and f_L with a probable error of 10 per cent.) for velocities up to double the velocity of sound. Rough values for a third coefficient f_H are also provided. It appears probable that the other components (except of course the drag) are much less important, and that values of the yaw up to perhaps 10 degrees may be regarded as small in this connection.

It is convenient to summarize here what we do and do not know about the components of the force system on the shells used in this trial. The values of the drag coefficient f_R may be regarded as known for all velocities at zero yaw. The values of f_M and f_L are roughly known for velocities up to $v/a = 2 \cdot 0$, and values of the yaw less than 10 degrees. From wind channel experiments f_R, f_M and f_L are all known for all values of the yaw when v/a is small, and these determinations probably apply so long as $v/a < 0 \cdot 7$. The damping effects are only known roughly, but sufficient is known to estimate how long a shell will take effectively to settle down to a steady state of motion.

On the other hand the variation of f_R with yaw is entirely unknown except from wind channel experiments, and so is the variation of f_M and f_L at values of the yaw greater than 10 degrees. The rate of diminution of the axial spin is unknown and so is the size of the swerve effect, though this latter is not likely to be important.

The variation of f_R with yaw could be studied experimentally by a suitable combination of jump card observations, with the use of the solenoid chronograph to determine as exactly as possible the deceleration of the shell at every point. The values of f_M and f_L for larger values of the yaw could be obtained by a detailed analysis of unstable rounds in which large values of the yaw are realized. A start could be made with the data of the present trial, but we cannot undertake this in this paper.

---

[*] The equation is now of the first order in η only, so that the exact solution may be written down in the form of an integral. By successive integration by parts we obtain the expansion (3.632) together with an integral representing the error after n terms.

In Part III., we have arrived at two separate solutions of the equations of motion of a shell treated as a rigid body, which together cover practically all types of motion which are likely to occur in practical shooting. (We ignore here the case of an unstable shell, since it is of no practical use.) A general solution of the equations of motion of type α has been developed, which applies with sufficient accuracy to the most general type of motion of a shell whose angle of yaw δ and inclination of the tangents of true and plane trajectories do not exceed (say) 0·1 radian. The solution of the equations of type β can be applied with sufficient accuracy to the steady (non-oscillatory) motion of a shell at any angle of yaw. In practice the large angles of yaw ($>$ 0·1 radian) only occur in the neighbourhood of or beyond the vertex of a high angle trajectory, and by this stage the initial angular oscillations of the shell have been completely damped out so that the condition for the applicability of the solution of type β is satisfied. Thus the solutions we have obtained, though theoretically inadequate, are probably sufficient to cover all cases likely to occur in practice.

In order to make use of these solutions to determine the complete motion of a shell, information is necessary as to the complete force system acting on the shell. Our information, as we have seen, is fairly complete for angles of yaw up to 10 degrees, and can be applied to calculate the true trajectory of any shell for which the angle of yaw does not exceed this value, if the loss of spin and increase of drag with yaw can be ignored.

Larger angles of yaw (exceeding 10 degrees) occur in general only as a consequence of the low velocity of the shell near the vertex of a high angle trajectory. The force system is then mainly covered by wind channel observations. The information as to the force system obtained by our methods is thus adequate for the calculation of a complete set of twisted trajectories at all elevations, at any rate for a 3-inch shell.

§ 5.1. *Problems for Further Discussion.*

5.11. *Unstable Rounds.*—We have already mentioned that further information about f_M and f_L, at yaws greater than 10 degrees, could be obtained by analysis of the unstable rounds. This requires the application of the exact top equation with a variable value of μ (§ 3.7) to the discussion. No means of introducing damping effects into these equations has yet been devised. It should, however, be possible to obtain fairly reliable information as to the variation of f_M and f_L with yaw between the angles of 10 degrees and 30 degrees by the analysis of the unstable rounds fired in this trial (Table IV.).

5.12. *Initial Conditions.*—By extrapolating the δ-curve and ϕ-curve backwards to the gun muzzle ($t = 0$) information may be obtained as to the way in which the projectile leaves the gun, which may prove of value. Owing to the effect of the initial oscillations on the ranging of the shell, it is important to determine whether, in general, the initial disturbance takes place at, or nearly at, the same orientation. Secondly, it is important to determine whether the initial oscillations may be regarded as caused by an impulsive couple whose size is independent of the twist of the rifling. If this is so, the amplitude of the initial oscillations of a shell can be cut down indefinitely by sufficiently increasing the spin. If, however, as appears to be the case from a rough survey of the data of the present trial, the initial circumstances are such that the impulsive couple (or its equivalent) increases in proportion to the twist of the rifling, then no increase of spin will reduce the oscillation below a certain definite limit. This conclusion would be technically important, as in the later stages of flight the spin is always in excess of requirements, and so the initial spin should be kept down as much as possible.

5.13. *Wind Effects.*—In calculating the effect of wind on a shell it is usual to assume that the shell at once turns its nose to the relative wind. This is not strictly correct, and the true angular motion in a wind when the velocity is known at every point can be determined by our theory, since the forces acting on the shell depend only on its motion *relative* to the air. Consider, for example, the special case in which a shell suddenly enters a cross-wind region from a region of still air; it starts its relative trajectory with a yaw δ given by the equation

$$\tan \delta = w/v,$$

where w is the wind velocity and v the velocity of the shell. At the same time $\delta' = 0$ and $\phi' = 0$. The equations of § 3.6 enable the subsequent motion to be properly traced, and the errors in the usual treatment calculated.

§ 5.2. *Effect of Size and Shape of Shell.*

The jump card trials described in this paper were carried out with shells of two different shapes only. The differences between the two shells may be seen from fig. 6 to be considerable, form A having an ogival head of roughly 2 calibres radius, while form B is of 6 calibres radius. For form B the experiments determine the moment coefficient only, for a single position of the centre of gravity, and give no information as to the cross-wind force. As experiments of this type are expensive and laborious to carry out, it is of importance to examine how far these results may be applied to shells of other shapes and sizes.

From the results of § 1.1 it appears that there is no evidence that the size (represented by the radius r of the shell) enters into any of the factors on which the force coefficients depend, so that the coefficients f_R, f_M, f_L may be considered as entirely independent of size. It is therefore sufficient to make experiments on shells of as small a calibre as is consistent with obtaining accurate measurements of the jump cards.

With regard to the effect of variation of shape we have very little evidence.

If we compare the moment coefficients f_M for shells of 2 and 6 calibres radius of head, as shown in figs. 4 and 5, it is obvious that the difference is much less marked than the difference between the two curves of f_R, and that the two curves of f_M are very nearly of the same shape. No great errors would be introduced by assuming that the values of f_M for the two shells were in a constant ratio. Thus it seems reasonable, until the appearance of evidence to the contrary, to consider that the value of the moment coefficient for any shell can be obtained by multiplying the value, obtained in this experiment, by a constant independent of the velocity. It will then be sufficient to determine the value of this constant at a single velocity, which may even be a low velocity attainable in a wind channel. The value of the cross-wind force factor for any shell may be obtained in a similar manner but the results will be more uncertain. For rough purposes it may even be sufficient to assume f_L and f_M independent of the velocity except when dealing with velocities very near the velocity of sound. It thus appears to be possible to treat f_M in the classical way in which f_R was treated, in which it was assumed that the values of f_R for two different shells are in a constant ratio at all velocities. This treatment is inadequate in the case of f_R, but on present evidence is far more valid in the case of f_M.

By applying the results of the present trial in this way, we may even hope to get reasonably accurate estimates of the drift and stability for any type of shell, on the basis of wind channel experiments only on the particular shape of shell required. The method would be especially valuable in connection with the design of new shapes of shell. It is known that, in general, the longer and more pointed a shell is, the less is its drag coefficient; by a series of wind channel tests on a series of shell shapes it would be possible to determine the greatest length of shell that would be sufficiently stable in a gun of given rifling, or the sharpness of rifling required to make a given shell stable. Useful information was obtained on this point from wind channel experiments before the jump card trial provided certain data for the extrapolation to high velocities. It must be emphasised, however, that this one experiment needs extension and confirmation before the structure sketched above can confidently be reared upon it.

We have now discussed in general terms the applicability of our theory and experiments to the calculation of drift, stability, the effect of wind, and the design of improved forms of shell. Though the details of the calculations on these various points are not given here, enough has been said to show that the results form some advance in the subject of the application of aerodynamics to the flight of shells.

TABLE V.—General Table of Results.

Column 1. Number of round.
„ 2. Muzzle velocity, f.s., for round, or mean for group.
„ 3. Air density ρ, lb./(ft.)^3 and temperature ° F.
„ 4. $\phi' = \frac{1}{2}\Omega$, degrees/sec.
Upper entry—Calculated value for round, or mean for group.
Lower entry—Observed value.[*]
„ 5. Period T, between successive minima of δ, in units of 10^{-3} sec.
„ 6. Maxima of yaw, $\alpha_1(t)$ degrees.
„ 7. Mean values of ΩT, radians, for each round, for mean velocity as stated.
„ 8. Minima[†] of yaw, $\beta_1(t)$ degrees.

[*] When there is only one entry there was no detectable difference between the observed and calculated values of ϕ.

[†] *Note.*—The sign given is the sign of $\beta_1(t)$ at the minimum, see § 4.0. The yaw δ is always positive.

Gun rifled 1 turn in 40 calibres.

1.	2.	3.	4.	5.	6.	7.	8.
				Type I.			
I. 11	918	0·0792 43°	2108	252 267 —	2·1 2·5 2·1?	19·10 905	0·0 —
I. 12	918	0·0792 43°	2108	276 262 —	6·1 6·9 7·5+	19·79 905	-0·4 -0·6
I. 14	920	0·0807 42°	2113	296 289	7·2 6·4	21·60 905	-0·9 -1·5?
I. 13	931	0·0807 42°	2139	234 260 —	1·5 3·1 5·5+	18·44 919	-1·0 -2·4?
I. 5	1565	0·0782 45°	3595	230 —	12·2 8·5	28·87 1539	-0·9
I. 6	1565	0·0782 45°	3595	227 —	13·7 9·5	28·49 1540	0·0

TABLE V. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Type I. (continued).							
I. 7	1565	0·0782 45°	3595 3375	354	10·5	41·71 1526	−2·3
I. 15	2130	0·0807 42°	4892	2T = 213	— ?7·5 5·5	18·19 2082	— −0·4
I. 16	2130	0·0807 42°	4892	2T = 203	— ?7·5 5·7	17·33 2085	— −0·6
I. 1	2167	0·0786 42°	4977	93½ 92 90½	4·5 4·1 3·7	15·98 2104	−1·0 −0·6 −0·6
I. 2	2167	0·0786 42°	4977	112 105 —	7·5 6·3 5·3	18·85 2117	0·0 −0·2 ?
I. 3	2167	0·0786 42°	4977	107 99½ —	4·5 3·5 3·1	17·94 2120	−0·3 −0·4
I. 4	2167	0·0786 42°	4977	118 121½	4·1 3·0	20·81 2113	−0·4 −1·4 ?
I. 19	2272	0·0812 40°	5217	98½ 90½ —	5·0 4·3 3·0	17·20 2217	−0·2 −0·3 ?
I. 20	2346	0·0812 40°	5388 5250	96½ 94½ —	8·7 7·9 7·1	17·96 2288	+0·4 +1·0
I. 21	2346	0·0812 40°	5388	99 99 —	5·6 4·0 3·3	18·03 2282	−1·0 −0·6

TABLE V. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Type II. Form A. C.G. forward.							
II. 8	934	0·0807 42°	1960 2020	210 224 —	1·6 3·2 3·8	15·30 923	−0·9 −0·5
II. 9	934	0·0807 42°	1960	246 228 —	4·5 4·0 4·7	16·22 922	−0·3 −1·2
II. 10	934	0·0807 42°	1960	254 256 —	6·7 6·8 8·1	17·45 921	−0·8 −0·5
II. 5	1585	0·0780 46°	3541	147½ 145½ —	4·9 3·7 2·4	18·11 1554	−0·7 −0·7
II. 6	1585	0·0780 46°	3541	144½ 145½ —	1·2 1·2 1·0	17·92 1555	−0·5? −0·5?
II. 7	1585	0·0780 46°	3541	187 170	3·1 2·2	22·07 1548	−0·4 −0·3?
II. 1	2024	0·0786 42°	4795 4530	99 89½ 91½	3·2 2·5 2·3	14·76 1983	−0·7 −0·8 −0·7
II. 2	2024	0·0786 42°	4795 4625	103 95 87	3·0 2·3 1·9	15·34 1982	0·0 −0·1 −0·5
II. 3	2024	0·0786 42°	4795 4435	96 89 92	2·9 2·6 2·3	14·29 1984	−0·2 −0·2 −0·2
II. 4	2024	0·0786 42°	4795	98½ 81 85	1·9 1·7 1·8	14·75 1985	0·0 0·0 0·0

TABLE V. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Type III. Form A. C.G. back.							
III. 8	931	0·0807 42°	2450	247 226 —	5·7 5·5 5·8	20·23 919	−0·4 −1·5
III. 9	931	0·0807 42°	2450	254 216 —	5·0 5·5 4·3	20·10 919	−0·4 −1·3
III. 10	931	0·0807 42°	2450	232 224 —	5·2 5·5 7·8	19·50 920	−0·3 −1·5
III. 5	1583	0·0780 46°	4166 4100	217 —	8·7 7·8	31·06 1556	−3·7
III. 6	1583	0·0780 46°	4166	196 —	9·0 8·2	28·52 1558	−3·6
III. 7	1583	0·0780 46°	4166 3980	237 —	5·5 6·3	32·93 1553	−2·9
III. 1	2025	0·0785 43°	5331	125 97 —	1·5 1·2 0·8	20·66 1994	— —
III. 2	2025	0·0785 43°	5331	113 97 —	3·7 3·4 3·4	19·54 1995	−0·7 −1·5
III. 3	2025	0·0785 43°	5331	109 107 —	2·4 1·8 1·3	20·10 1994	−0·2? −0·5
III. 4	2025	0·0785 43°	5331	106 102 —	2·9 2·1 1·8	19·36 1994	0·0 −0·8

TABLE V. (continued).

Gun rifled 1 turn in 30 calibres.

1.	2.	3.	4.	5.	6.	7.	8.
			Type I. Form A. Empty shell.				
I. 22	1119	0·0811	3429	91½	2·1	10·07	+0·1
				86½	1·9		0·0
				82	1·9		−0·3?
		38°		86	—		—
				79½	1·9		−0·2
				87	2·0	1089	—
I. 23	1119	0·0811	3429	79	3·0	10·34	+0·4
				91	2·4		−0·1
				87	2·2		−0·4
		38°		86	—		—
				83½	2·0		−0·4
				84½	1·8	1090	—
I. 24	1119	0·0811	3429	85	2·8	10·27	+0·4
				90	2·4		−0·2
				86	2·2		−0·4
		38°		84½	—		—
				84½	2·1		−0·6
				84	2·1	1090	−0·9?
I. 25	1326	0·0811	4061	74½	2·0	9·91	0·0
				71	1·8		0·0
				69	1·9		0·0
		38°		68½	—		—
				71	1·6		0·0
				70	1·4	1280	−0·8?
				—	2·0		
I. 26	1326	0·0811	4061	71	3·6	10·04	+0·4
				74	2·8		+0·1
				71	2·6		0·0
		38°		73	—		—
				66½	2·6		−0·1
				69½	2·2	1285	−0·8?
I. 27	1563	0·0805	4786	60½	3·8	9·59	−0·3
				55½	3·7		−0·3?
				59	3·5		−0·2
		36°		60½	—		—
				55½	2·9		−0·5
				56½	3·3?		−0·8
				—	2·5	1515	

TABLE V. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Type I. Form A. Empty shell (continued).							
I. 28	1563	0·0805	4786	59	3·1	9·66	−0·4
				59	2·6		−0·4
				59	2·1		−0·3
		36°		59	—		—
				57½	1·9		−0·2
				54½	2·0		−0·5 ?
				—	2·1	1515	
Type II. Form A. C.G. forward.							
II. 17	1119	0·0811	3168	92½	2·8	9·85	0·0
				90	2·75		0·0
				88½	2·6		−0·6
		38°	3128	91	—		−0·5
				91½	2·0		−0·7
				—	1·7	1092	
II. 18	1119	0·0811	3168	88	3·3	9·88	+0·1
				94½	2·8		0·0
				87½	2·6		−0·6
		38°	3128	89	—		−0·7
				91	2·3		−0·7
				—	2·2	1093	
II. 19	1119	0·0811	3168	88	2·9	9·82	+0·1
				89	2·4		0·0
				89½	2·0		−0·3
		38°	3128	90	—		−0·4
				89½	2·0		−0·5 ?
				91½	2·1	1088	−0·5 ?
II. 24	1292	0·0805	3709	76	2·7	9·70	−0·15
				74	2·9		−0·3
				72	2·7		−0·2
		36°		76	—		−0·3
				75	2·4		−0·7
				76	3·0	1259	−0·5
II. 22	1589	0·0819	4738	52	2·1	8·93	+0·5
				51	1·9		+0·1 ?
				54½	1·4		—
		36°		56½	1·4		—
				54½	—		+0·5
				52	1·1		—
				55½	1·4	1543	+0·1

Table V. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Type II. Form A. C.G. forward (continued).							
II. 23	1589	0·0819	4738	54	2·3	9·13	+0·3
				5 periods	2·1		+0·1
				in 276	—		—
		36°			—		—
					—		—
					1·9		—
					1·6	1548	-0·1
Type III. Form A. C.G. back.							
III. 17	1119	0·0811	3928	67	1·8	10·08	0·0
				75	1·8		0·0
				71	2·2		0·0
		38°		73	2·2		0·0
				71	—		0·0
				74	1·8		0·0
				77	1·7	1091	0·0
III. 18	1119	0·0811	3928	63	1·4	9·91	+0·2
				71	1·4		—
				75	1·4 ?		+0·1
		38°		77	1·8		—
				72	—		+0·2
				71½	1·4		+0·1
				67	1·3	1091	-0·1
III. 19	1119	0·0811	3928	64	1·6	10·17	+0·1
				72	1·4		+0·1
				77	1·3		+0·1
		38°		74	1·4		+0·1
				74	—		+0·1
				71	1·1		-0·1
				77	1·1	1091	—
III. 20	1292	0·0805	4534	62	2·7	9·63	0·0
				59	2·7		-0·1
				62	2·2		-0·5
		36°		58	2·2		—
				63	—		-1·0
				63	2·0	1261	-0·5
				60	1·6		-0·9
				—	1·6		

TABLE V. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Type III. Form A. C.G. back (continued).							
III. 21	1292	0·0805	4534	60	3·6	9·50	0·0
				62	3·2		0·0
				59	2·7		−0·2 ?
		36°		59	2·6		—
				61	—		−0·7 ?
				59	2·3		−0·5
				60	2·1	1262	−0·6
III. 22	1567	0·0805	5501	42	3·5	9·38	−0·1
				50	2·9		−0·2
				49	2·8		−0·4
		36°		49	2·4		—
				50	—		−0·4
				49	2·6		−0·3
				46	—		−0·3
				—	2·3	1525	
III. 23	1567	0·0805	5501	51 ?	1·7 ?	9·22	0·0
				42	1·6		0·0
				51	1·4		0·0
		36°		47	1·0		—
				48 ?	—		—
				51 ?	1·0 ?		0·0
				49 ?	—	1526	0·0
					1·0		0·0
Type IV. Form B.							
IV. 21	900	0·0811	2431	116	1·7	10·27	−0·6
				120	2·8		−1·5
				122	2·3		—
		38°		116	—		−1·0
				126	2·7 +	884	−2·0 ?
				—	3·5 +		
IV. 22	900	0·0811	2431	124	3·0	10·10	+0·3
				128	2·8		−0·2
				119	1·9		—
		38°		110	—		−0·4
				119	2·3	884	−0·7
				—	3·0 +		

TABLE V. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Type IV. Form B (continued).							
IV. 23	900	0·0811 38°	2431	— 112 120 114 118 —	1·7? 1·7 2·6 — 2·0 2·0+	9·84 884	+0·2 −0·8 — −0·8 −1·0
IV. 13	1078	0·0811 38°	2911	113 128 125 113 —	2·0 1·8 1·5? 1·6 2·0	12·39 1059	+0·3 −0·4 −0·6 −0·9
IV. 14	1078	0·0811 38°	2911	109 123 110 105 —	1·6 1·5 1·8? 1·6 1·7	11·45 1060	+0·2 −0·3 — −0·3
IV. 15	1078	0·0811 38°	2911	107 121 118 108 —	2·2 2·1 2·2 1·7 1·8	11·76 1060	+0·3 −0·1 — −0·5
IV. 16	1547	0·0811 38°	4178	82 78 67 75 74?	1·6 1·6 1·4 1·0 0·9	10·72 1503	+0·1 0·0 — 0·0 −0·1
IV. 17	1547	0·0811 38°	4178	83 75 75 73 71	3·5 3·2 2·8 1·9 2·6	10·72 1503	+0·1 0·0? — −1·2? −0·9?
IV. 18	1547	0·0811 38°	4178	83 80 75 73 —	4·2 3·9 3·3 3·2 3·1	11·09 1509	+0·1 0·0 — −0·4 −1·1

Table V. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Type IV. Form B (continued).							
IV. 19	1547	0·0811 38°	4178	77 75 73	4·0 3·4 3·4	10·96 1496	—
IV. 20	1547	0·0811 38°	4178	73 75 72	3·3 2·5 2·2	10·69 1497	—
IV. 24	2101	0·0805 36°	5675	55 57 51 55 53	4·6 3·8 3·5 3·5 3·4	10·70 2045	−0·2 −0·3 — −0·5 ? −1·0 ?
IV. 25	2112	0·0805 36°	5705	56 55 52 55 54	2·2 2·2 1·3 1·2 1·3	10·70 2073	0·0 ? 0·0 — 0·0 ? −0·8
IV. 26	2149	0·0805 36°	5805	50 53 54 53 54	2·4 2·0 2·0 1·7 1·2	10·60 2093	−0·1 0·0 — — −0·3

TABLE VI.—Values of the Stability Coefficient and the Air Couple deduced from Analysis of the *Stable* Shells.

Summary of Notation used in the Headings of this Table.

$\mu \sin \delta$ = couple due to air forces.

$s = \dfrac{\Omega^2 B}{4\mu}$ = stability coefficient.

v = mean velocity of shell, f.s.

ρ = air density, lb./(ft.)^3.

a = velocity of sound, f.s.

$f_M(v/a) = \mu/(\rho v^2 r^3)$, the air couple coefficient.

N.B.—The values of s, μ and f_M have not been corrected in this table for the effect of the cards.

1. Round No. or group of rounds.	2. Twist of rifling.	3. Value of s deduced from observation.	4. Mean value of v corresponding to value of s.	5. Value of μ.	6. Value of $f_M(v/a)$.	7. Value of v/a.	8. Total percentage spread of s or μ in group.
I. 11, 12	1/40	1·113	905	1220	9·58	0·824	4·1
I. 14	1/40	1·087	906	1250	9·71	0·825	
I. 13	1/40	1·131	919	1230	9·26	0·837	
I. 22–24	1/30	1·61	1090	2230	11·85	0·996	3·2
I. 25, 26	1/30	1·66	1283	3030	11·62	1·173	1·7
I. 27, 28	1/30	1·74	1515	4000	11·09	1·388	1·1
I. 5– 7	1/40	1·005	1535	3920	10·89	1·394	0·9
I. 15, 16	1/40	1·137	2084	6410	9·36	1·897	6·7
I. 1	1/40	1·180	2104	6390	9·40	1·916	
I. 2– 4	1/40	1·118	2117	6750	9·80	1·927	
I. 19	1/40	1·152	2217	7190	9·22	2·023	1·9
I. 20, 21	1/40	1·133	2285	7800	9·41	2·085	
II. 8–10	1/40	1·172	922	1170	8·71	0·840	4·9
II. 17–19	1/30	1·69	1091	2000	10·46	0·997	0·8
II. 24	1/30	1·72	1259	2700	10·86	1·153	
II. 5– 7	1/40	1·121	1553	3680	9·98	1·408	4·7
II. 22, 23	1/30	1·94	1546	3800	9·95	1·416	3·8

TABLE VI. (continued).

1.	2.	3.	4.	5.	6.	7.	8.
Round No. or group of rounds.	Twist of rifling.	Value of s deduced from observation.	Mean value of v corresponding to value of s.	Value of μ.	Value of $f_{\rm M}(v/a)$.	Value of v/a.	Total percentage spread of s or μ in group.
II. 1	1/40	1·216	1983	5506	9·12	1·804	
II. 2	1/40	1·200	1982	5689	9·42	1·803	
II. 3	1/40	1·232	1984	5320	8·80	1·805	
II. 4	1/40	1·220	1985	5809	9·59	1·806	
III. 8–10	1/40	1·107	919	1520	11·41	0·837	0·9
III. 17–19	1/30	1·64	1091	2570	13·66	0·997	3·3
III. 20, 21	1/30	1·76	1262	3200	12·79	1·156	2·1
III. 22, 23	1/30	1·84	1526	4500	12·32	1·398	2·9
III. 5– 7	1/40	1·035	1556	4410	11·96	1·411	0·6
III. 1– 4	1/40	1·109	1994	7020	11·52	1·814	1·4
IV. 21–23	1/30	1·64	884	1270	10·25	0·808	2·6
IV. 13–15	1/30	1·39	1060	2140	12·06	0·969	5·0
IV. 16–20	1/30	1·505	1502	4070	11·41	1·373	3·7
IV. 24	1/30	1·525	2045	7420	11·29	1·874	1·0
IV. 25	1/30	1·525	2075	7500	11·11	1·899	
IV. 26	1/30	1·54	2093	7680	11·16	1·917	

Table VII.—Observed Values of $h+\kappa$ and $h-\kappa+2\gamma$, the Damping Factors for each Group.

Groups fired at a velocity of 900 f.s. apparently have negative damping and are not included.

Group. Muzzle velocity, f.s.	Calculated value of κ.	$h+\kappa$.	$h-\kappa+2\gamma$.	Probable value of h.
I. 22–24 1119	0·4	1·9	1·6 ?	1·8
I. 25, 26 1326	0·3	2·4	1·2	1·8
I. 27, 28 1563	0·4	3·0	0·5	1·8
I. 1– 4 2167	0·7	2·2	0·6	1·4
I. 19–21 2320	0·8	2·2	−0·2 ?	1·3
II. 17–19 1119	0·4	2·2	1·2	1·7
II. 24 1292	0·2	0·9	0·6	1·5
II. 5– 7 1585	0·4	3·4	0·4	2·0
II. 22, 23 1589	0·4	3·3	0·6	
II. 1– 4 2027	0·6	3·0	0·6	2·0
III. 17–19 1119	0·4	0·7 ?	0·1 ?	1·0
III. 20, 21 1292	0·2	3·1	0·9	2·0
III. 22, 23 1567	0·4	3·0	0·3	2·0
III. 1– 4 2025	0·6	4·2	1·7	3·0
IV. 13–15 1078	0·5	0·7	1·4	1·0
IV. 16–18 1547	0·5	3·1	1·2	2·0
IV. 24–26 2120	0·7	5·0	0·9	3·0

N.B.—The calculated value of κ is obtained by using the value of the cross-wind force coefficient given in Table I.

TABLE VIII.—Plane Trajectories at 50 degrees and 30 degrees, with Calculations of the Drift, &c., for Shells of External Form A.

Constants used in the Calculations.

Muzzle velocity		2000 feet per second.
Centre of gravity		4·88 inches from the base.
Weight		16·02 lb.
Moments of inertia	A	0·1329 lb. (ft.)².
	B	1·1555 lb. (ft.)².
Ω for	gun rifled 1 in 30	192·7 radians per second.
	„ 1 „ 40	144·5 „ „
Ballistic coefficient		1·75.

TABLE VIIIA.—Plane Trajectories at Elevations of 50 degrees and 30 degrees.

Column 1. Time t, seconds.
„ 2. Velocity v_1, feet per second.
„ 3. Inclination θ_1, degrees.
„ 4. Horizontal distance X, feet.
„ 5. Vertical height Y, feet.

Elevation 50 degrees.					Elevation 30 degrees.			
1.	2.	3.	4.	5.	2.	3.	4.	5.
		° ′				° ′		
0	2000	50 0	0	0	2000	30 0	0	0
1	1720	49 21	1,199	1,413	1726	29 8	1,614	917
2	1506	48 36	2,254	2,628	1515	28 8	3,033	1692
3	1342	47 44	3,201	3,686	1352	26 59	4,300	2354
4	1218	46 45	4,068	4,624	1230	25 42	5,454	2926
6	1059	44 28	5,648	6,241	1075	22 45	7,538	3867
8	959	41 48	7,116	7,619	985	19 24	9,455	4608
10	877	38 43	8,514	8,805	916	15 42	11,264	5182
12	807	35 13	9,857	9,819	860	11 40	12,988	5603
14	746	31 15	11,155	10,671	816	7 19	14,639	5881
16	693	26 45	12,411	11,369	780	2 43	16,227	6021
18	647	21 44	13,631	11,921	751	− 2 7	17,755	6030
20	609	16 10	14,817	12,330	728	− 7 5	19,228	5912
24	557	3 40	17,098	12,738	703	− 17 11	22,017	5314
28	537	− 9 50	19,266	12,624	700	− 26 55	24,608	4261
32	547	− 22 55	21,331	12,013	712	− 35 51	27,011	2790
36	579	− 34 25	23,293	10,930	735	− 43 40	29,228	938
40	627	− 43 54	25,151	9,403				
44	681	− 51 30	26,902	7,466				
48	735	− 57 32	28,539	5,157				
52	786	− 62 23	30,057	2,521				
55	818	− 65 25	31,114	359				

TABLE VIIIB.—Calculation of the Drift, Stability, and Damping Factors, for the Gun Rifled 1 in 30.

Column 1. The time t, seconds.
„ 2. The stability factor s.
„ 3. $-4s\theta'_1/\Omega$.
„ 4. $4s\kappa/\Omega$.
„ 5. ψ, equation (4.203).
„ 6. The drift Z, feet.
„ 7. The azimuth, arc tan (Z/X), degrees.
„ 8. The second term in the expansion of $\bar{\eta}$, equation (3.632), given by

$$\frac{\eta^{(2)}}{(i\Omega)^2} = \frac{4s}{(i\Omega)^2}\left\{\frac{d}{dt}(4s\theta'_1) + 4s\kappa\theta'_1 + c'\int_0^t 4s\kappa\theta'_1\,dt/c\right\}.$$

„ 9. The damping factor (§ 4.22),

$$(\sigma_0/\sigma)^{\frac{1}{2}}\,e^{-(q_1-q_2)}.$$

Elevation 50 degrees.

1.	2.	3.	4.	5.	6.	7.	8.	9.
0	1·945	0·00042	0·0217	0	0			1
1	2·598	0·00062		0·00038	0·2			0·702
2	3·334	0·00098	0·0214	0·00082	0·8			0·520
3	4·189	0·00139		0·00129	1·8			0·404
4	5·129	0·00192	0·0221	0·00182	3·1		0·00007	0·330
6	7·123	0·00321		0·00405	7·5	0 5		0·196
8	10·95	0·00569	0·0570	0·00752	15·9		0·00066	0·115
10	14·60	0·00868		0·0114	29·1	0 12		0·076
12	18·39	0·01240	0·0518	0·0155	47·3	0 16	0·00155	0·055
14	22·20	0·01693		0·0199	70·4			0·043
16	26·33	0·0227	0·0585	0·0250	98·6	0 27	0·00334	0·034
18	30·88	0·0297		0·0308	133			
20	35·16	0·0372	0·0663	0·0374	173		0·0056	
24	42·63	0·0511		0·0530	276	0 55		
28	45·87	0·0562	0·0755	0·0705	409	1 13		
32	43·24	0·0486		0·0885	573	1 32		
36	37·11	0·0353	0·0700	0·1050	762		0·0012	
40	30·11	0·0231		0·1192	971	2 12		
44	24·02	0·0147	0·0595	0·1315	1190			
48				0·1420	1415		−0·0019	
52				0·1510	1637			

TABLE VIIIB. (continued).

Elevation 30 degrees.

1.	2.	3.	4.	5.	6.	7.
						° '
0	1·945	0·00056		0	0	
1	2·55	0·00086				
2	3·21	0·00122		0·00080	1·1	
4	4·79	0·00234		0·00174	4·1	
6	8·43	0·00368		0·00351	9·5	
8	9·11	0·00581		0·00655	19·2	0 7
10	11·16	0·00785		0·01022	34·4	
12	14·01	0·01063		0·01417	55·3	0 15
14	15·75	0·01276		0·01828	82·0	
16	17·48	0·01495		0·0225	114·3	0 24
18	18·85	0·01671		0·0269	152·0	
20	19·92	0·01810		0·0316	195·0	0 35
24		0·0191		0·0418	297	0 46
28		0·0174		0·0525	419	0 59
32		0·0143		0·0629	557	1 11
36		0·0110		0·0729	708	1 23

TABLE IX.—Comparison of Calculated Drift with Observations of April–May and February, 1918.

The azimuth of the shell at time t (in minutes of angle) = At.

Elevation 50 degrees.

Observations of April–May.

Rifling 1/30.			Rifling 1/40.		
Mean observed time.	Mean observed A.	Calculated A.	Mean observed time.	Mean observed A.	Calculated A.
10·9	1·46	1·27	10·2	1·18	0·90
23·9	2·24	2·29	22·9	1·30	1·66
33·3	2·89	2·85	31·0	1·95	2·10
41·3	3·16	3·36	39·1	2·00	2·44

TABLE IX. (continued).

Elevation 50 degrees (continued).

Observations of February.

Rifling 1/30.			Rifling 1/40.		
Mean observed time.	Mean observed A.	Calculated A.	Mean observed time.	Mean observed A.	Calculated A.
6·99	4·54	0·86	6·33	3·61	0·63
15·03	1·68	1·65	14·07	1·40	1·18
26·08	1·99	2·46	24·93	1·47	1·86

Elevation 30 degrees.

Observations of April–May.

Rifling 1/30.			Rifling 1/40.		
Mean observed time.	Mean observed A.	Calculated A.	Mean observed time.	Mean observed A.	Calculated A.
10·04	0·975	1·05	9·58	1·63	0·79
20·6	1·575	1·77	19·35	0·80	1·28
27·9	2·09	2·06	25·95	1·04	1·50

Observations of February.

Rifling 1/30.			Rifling 1/40.		
Mean observed time.	Mean observed A.	Calculated A.	Mean observed time.	Mean observed A.	Calculated A.
13·2	1·40	1·32	13·02	1·73	0·97
22·52	1·36	1·86	22·05	1·25	1·38